



(19)

(11) Publication number:

10336202 A

Generated Document.

PATENT ABSTRACTS OF JAPAN

(21) Application number: 09144791

(51) Intl. Cl.: H04L 12/28 G06F 13/00 H04L 29/08 H04Q 3/00

(22) Application date: 03.06.97

(30) Priority:

(43) Date of application
publication: 18.12.98(84) Designated
contracting states:

(71) Applicant: FUJI XEROX CO LTD

(72) Inventor: YOSHIMURA KOICHI
UESAWA ISAO
MITSUTAKE KATSUYA

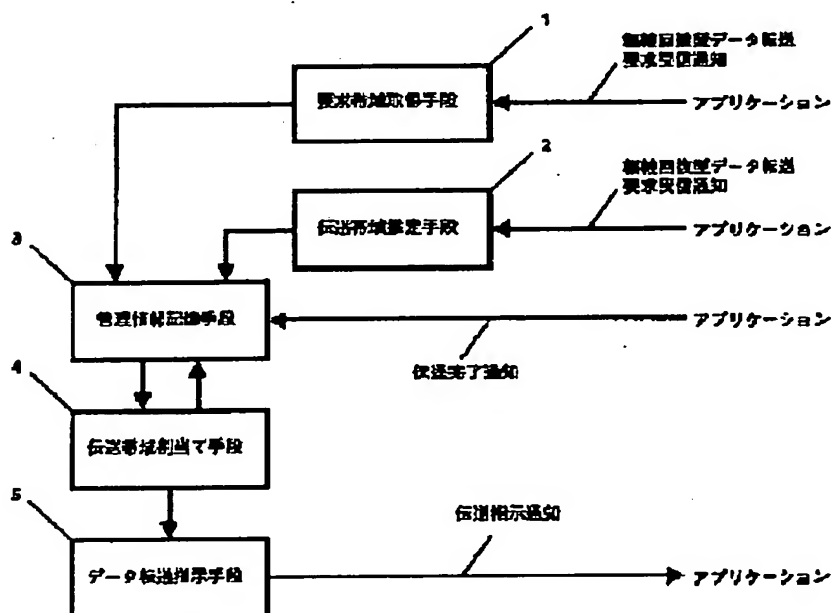
(74) Representative:

(54) DATA TRANSFER DEVICE
AND ITS METHOD

(57) Abstract:

PROBLEM TO BE SOLVED: To conduct recovery type congestion control and avoidance type congestion control simultaneously.

SOLUTION: A request band acquisition means 1 acquires a request transmission band for congestion avoidance type data transfer based on control information exchanged between applications prior to data transfer. A transmission band estimate means 2 estimates a transmission band for congestion avoidance type data transfer based on the control information exchanged between applications prior to data transfer. A management information storage means 3 stores management information for all of congestion avoidance type data transfer and congestion recovery type data transfer that use the transmission channel band at present. A transmission band assignment means 4 discriminates the propriety of execution of data transfer based on the remaining transmission channel band and the requested transmission band for data transfer requested newly and assigns the transmission band to the congestion avoidance type data transfer requested newly. A data transfer instruction means 5 instructs transmission to the application as to new data transfer request.



COPYRIGHT: (C)1998,JPO

(51) Int.Cl. ⁸	識別記号	F I	
H 0 4 L 12/28		H 0 4 L 11/20	G
G 0 6 F 13/00	3 5 1	G 0 6 F 13/00	3 5 1 A
H 0 4 L 29/08		H 0 4 Q 3/00	
H 0 4 Q 3/00		H 0 4 L 13/00	3 0 7 C

審査請求 有 請求項の数13 O L (全 38 頁)

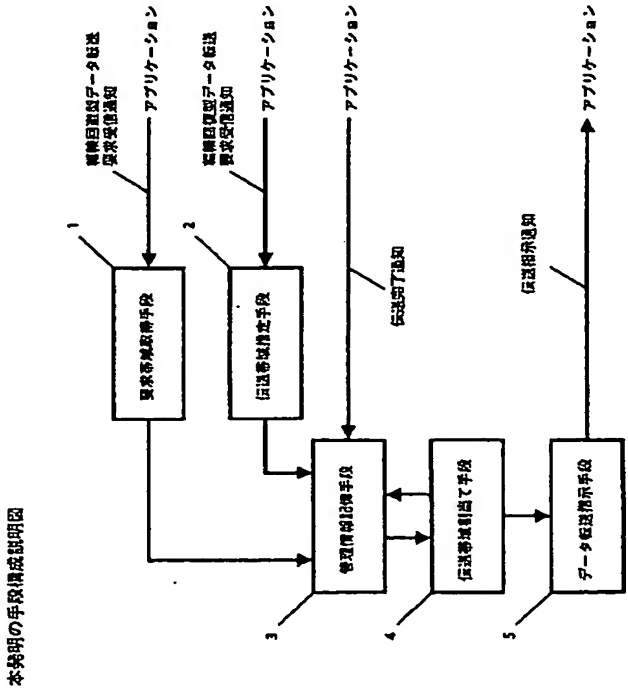
(21) 出願番号	特願平9-144791	(71) 出願人	000005496 富士ゼロックス株式会社 東京都港区赤坂二丁目17番22号
(22) 出願日	平成9年(1997)6月3日	(72) 発明者	▲吉▼村 浩一 神奈川県足柄上郡中井町境430 グリーン テクなかい 富士ゼロックス株式会社内
		(72) 発明者	上澤 功 神奈川県足柄上郡中井町境430 グリーン テクなかい 富士ゼロックス株式会社内
		(72) 発明者	光武 克也 神奈川県足柄上郡中井町境430 グリーン テクなかい 富士ゼロックス株式会社内
		(74) 代理人	弁理士 澤田 俊夫

(54) 【発明の名称】 データ転送装置および方法

(57) 【要約】

【課題】 回復型輻輳制御と回避型輻輳制御とを同時に行えるようにする。

【解決手段】 要求帯域取得手段1は、データ転送に先だってアプリケーション間で交換される制御情報から、輻輳回避型データ転送の要求伝送帯域を取得する。伝送帯域推定手段2は、データ転送に先だってアプリケーション間で交換される制御情報から、輻輳回復型データ転送の伝送帯域を推定する。管理情報記憶手段3は、現時点で伝送路帯域を使用中の輻輳回避型データ転送、輻輳回復型データ転送の全てについて、その管理情報を記憶する。伝送帯域割当て手段4は、伝送路帯域の未使用量と、新規に要求されたデータ転送の要求伝送帯域とから、前記データ転送の実行の可否を判定すると共に、新規要求の輻輳回避型データ転送に対して伝送帯域を割当てる。データ転送指示手段5は、新規のデータ転送要求について、アプリケーションに送信を指示する。



【特許請求の範囲】

【請求項1】 時分割交換ネットワークに接続された複数のコンピュータ上のアプリケーションが前記ネットワークを介してデータ転送を行い、特定のコンピュータに対して他の複数のコンピュータから単一の経路を介して同時にデータ転送が行われるネットワークシステムにおいて、

前記時分割交換ネットワークにおける輻輳を検出する輻輳検出手段と、

前記特定のコンピュータによる前記アプリケーションのデータ転送要求の受信を監視して前記時分割交換ネットワークの資源使用総量を判別する資源使用量観測手段と、

新規に要求されるデータ転送が第1の種類のデータ転送のときに、前記時分割交換ネットワーク上で、前記データ転送の要求に応じてデータ転送を開始し、前記輻輳検出のフィードバックにより、前記時分割交換ネットワークの資源使用量を制御する第1の転送制御手段と、

新規に要求されるデータ転送が第2の種類のデータ転送のときに、前記資源使用量観測手段で判別した資源使用量に基づき前記新規要求されるデータ転送に対する前記時分割ネットワークの資源使用可能量を求め、前記資源使用可能量の範囲で資源割当て量を通知する資源割り当て手段と、

前記資源割当て量の通知に基づいて、前記新規に要求された第2の種類のデータ転送を前記資源割当て量の帯域で実行する第2の転送制御手段とを有することを特徴とするデータ転送装置。

【請求項2】 前記資源使用量観測手段は、前記第1のデータ転送の転送パスおよび前記第2のデータ転送の転送パスが集中する位置において、前記複数のコンピュータと前記特定のコンピュータ間の接続手順のための制御パケットを監視して前記時分割交換ネットワークの資源使用総量を判別する請求項1記載のデータ転送装置。

【請求項3】 前記資源使用量観測手段は、前記第1のデータ転送の要求を受信した時点でこのデータ転送に要するネットワーク帯域の中心値とその時間変動量を推定して、前記第1のデータ転送による前記時分割交換ネットワークの資源使用量を集計し、さらに前記第2のデータ転送による前記時分割交換ネットワークの資源使用量を集計し、これら集計した値を加算して前記時分割交換ネットワークの資源使用総量を判別する請求項1または2記載のデータ転送装置。

【請求項4】 前記第1のデータ転送を、対応するアプリケーションの種別および対応するアプリケーションの処理パラメータの少なくとも1つから予め設定した複数のカテゴリーに分類し、前記カテゴリーに対応して予め設定した伝送帯域値を前記第1のデータ転送のネットワーク帯域の中心値と推定する請求項3記載のデータ転送装置。

【請求項5】 前記第1のデータ転送に要するネットワーク帯域の時間変動量は、前記前記資源使用量観測手段において集計したデータ転送の総量の短時間変動量観測値に、一定の係数を乗じることにより決定する請求項3または4記載のデータ転送装置。

【請求項6】 前記資源割り当て手段は、前記第2のデータ転送の許可時間間隔を一定の時間に分割し、この時間間隔ごとに資源割当てを実行し、前記資源使用量観測手段は一定のデータパケット到着時間間隔に基づいて前記データ転送の総量の短時間変動量観測値を算出する請求項5記載のデータ転送装置。

【請求項7】 前記他の複数のコンピュータはクライアントであって、前記特定のコンピュータはサーバーである請求項1、2、3、4、5または6記載のデータ転送装置。

【請求項8】 前記特定のコンピュータはプリントサーバーである請求項1、2、3、4、5、6または7記載のデータ転送装置。

【請求項9】 前記特定のコンピュータは前記ネットワーク内のスイッチである請求項1、2、3、4、5、6または7記載のデータ転送装置。

【請求項10】 前記特定のコンピュータは、前記データ転送の相手先を選択するネットワーク通信制御サーバーである請求項1、2、3、4、5、6または7記載のデータ転送装置。

【請求項11】 ネットワークにより接続された複数のコンピュータの間のデータ転送を行うデータ転送装置において、

前記ネットワークにおける輻輳を検出する輻輳検出手段と、

前記輻輳検出手段による輻輳検出をフィードバックして前記ネットワークの資源利用量を制御しながら前記コンピュータ間のデータ転送を行う第1の転送制御手段と、前記データ転送が集中する経路における前記ネットワークの資源使用量を観測する資源使用量観測手段と、

前記資源使用量観測手段で観測された資源使用量に基づいて、使用可能な範囲で前記ネットワークの資源を割り当てて、前記コンピュータ間のデータ転送を実行する第2の転送制御手段とを有することを特徴とするデータ転送装置。

【請求項12】 ネットワークにより接続された複数のコンピュータの間のデータ転送を行うデータ転送方法において、

前記ネットワークにおける輻輳を検出する輻輳検出ステップと、

前記輻輳検出ステップによる輻輳検出をフィードバックして前記ネットワークの資源利用量を制御しながら前記コンピュータ間のデータ転送を行う第1の転送制御ステップと、

前記データ転送が集中する経路における前記ネットワー

クの資源使用量を観測する資源使用量観測ステップと、前記資源使用量観測ステップで観測された資源使用量に基づいて、使用可能な範囲で前記ネットワークの資源を割り当てて、前記コンピュータ間のデータ転送を実行する第2の転送制御ステップとを有することを特徴とするデータ転送方法。

【請求項13】 ネットワークにより接続された複数のコンピュータの間のデータ転送を行うために用いられるデータ転送用コンピュータ・プログラム製品において、前記ネットワークにおける輻輳を検出する輻輳検出ステップと、
前記輻輳検出ステップによる輻輳検出をフィードバックして前記ネットワークの資源利用量を制御しながら前記コンピュータ間のデータ転送を行う第1の転送制御ステップと、
前記データ転送が集中する経路における前記ネットワークの資源使用量を観測する資源使用量観測ステップと、前記資源使用量観測ステップで観測された資源使用量に基づいて、使用可能な範囲で前記ネットワークの資源を割り当てて、前記コンピュータ間のデータ転送を実行する第2の転送制御ステップとを所定のコンピュータに実行させるために用いることを特徴とするデータ転送用コンピュータ・プログラム製品。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 この発明は、複数のコンピュータがネットワークを共有してデータ転送を行う環境に好適なデータ転送技術に関するものである。

【0002】

【従来の技術】 先ず、説明のために、本発明の対象とするコンピュータネットワークについて述べる。

【0003】 図1にコンピュータネットワークの概略構成図を示す。図中、複数のコンピュータ100が物理的に分散された場所に配置される。これらコンピュータ100は、お互いにネットワーク101を介して接続される。各々のコンピュータ100には、複数のアプリケーション（適応業務対応のプログラム）102が存在する。データ通信制御ソフトウェア103は、アプリケーション102の要求するデータ通信を所定の手順に従って実行する。ネットワーク101はこのデータ通信を実現する伝送媒体である。これによって、複数のコンピュータ100間で、アプリケーション102の要求するデータ転送を実現している。この構成は、一般にLAN（Local Area Network）として知られ、例えば、Ethernetなどで常套的に実現されている。

【0004】 ところで、このようなコンピュータネットワークにおいては、分散したコンピュータ100上のアプリケーション102が、実質的なデータ転送の主体である。アプリケーション102の機能から、以下のよう

にデータ転送の総量、速度が規定される。これを一般的に、トラヒックと呼んでいる。

①分散されたアプリケーション102間の制御

・分散したアプリケーション102間での情報交換が目的である。

・1回の転送データ量は数キロ・バイト（1キロは10の3乗）程度である。例えば、RPC（IETF RFC1057）の転送データ量は約1.5キロ・バイトである。

・伝送遅延に対する制約が高い。

②アプリケーション102の処理結果の転送

・主に蓄積されている処理結果データを要求しているコンピュータに転送することが目的である。

・1回の転送データ量は数百キロ・バイトから数百メガ・バイト（1メガは10の6乗）程度まで、データの内容や処理の総量によって異なる。例えば、A4サイズ、400dpi、白黒16階調画像のプリントでは、転送データ量は約8メガ・バイトである。

・伝送遅延に対する制約は、比較的少ない。一方、伝送媒体の速度に対する要求はデータ量に大きく依存する。

【0005】 いずれのトラヒックも、コンピュータ間通信という用途から、以下の要求条件を満足する必要がある。前記通信制御プログラム103は、これを実現するための手順をパケット単位に実行する。

1) 伝送誤りからの回復。すなわち送達確認手順の実行。

2) アプリケーション102が起動した時刻にデータ転送を即時に実行する。すなわち即時データ転送開始。

【0006】 一般にこれをコネクションレス通信と呼び、コンピュータ起動中は、常時ネットワークを使用可能な状態に維持する。

【0007】 一方、ネットワーク101は、データ転送を行うための手段と位置づけられる。伝送媒体は、伝送路とそれを接続するパケット交換スイッチとから構成される。一般に、このようなネットワーク101をスイッチ型パケット交換ネットワークと呼ぶ。前記パケット交換スイッチには、パケットが一時的に滞留するバッファ101a（図2参照）が伝送パス毎に存在する。

【0008】 つぎに、前記コンピュータネットワーク上での従来のデータ転送装置に関して説明する。図2は、従来のデータ転送装置の概略構成を示したものである。図中、輻輳検知手段104は、ネットワーク上101での後述する輻輳状態を検出し、送信側に輻輳状態であることを通知するフィードバック制御部である。送信側の輻輳回復型送信帯域規制手段105では、アプリケーション102によってデータ転送が要求された時点で、要求された送出速度、もしくは自らの最大能力の送出速度でデータ転送を開始する。輻輳検知手段104から輻輳状態を通知されている期間、パケットデータの送出を規制、又は、送出速度を低減させ、輻輳状態から復帰する

ための手順を実行する。

【0009】このような制御が必要となる理由について説明する。

【0010】各コンピュータ間のデータ転送は、任意の構成のネットワーク、即ち伝送媒体を一定時間占有する。この占有の程度を、ネットワーク資源の使用量と呼ぶ。具体的には、送信側と、受信側を結ぶ伝送媒体の伝送路帯域と継続時間で定義される。

【0011】LANではこの一定の伝送媒体（伝送パス）上に複数のコンピュータ、又は、複数のアプリケーションを同時に接続する。これにより、伝送媒体の規模、構成コストを最小化する。このことから、前記ネットワーク上では、複数の伝送要求が競合する事を前提とする制御が必要となる。

【0012】伝送媒体が提供しうるネットワーク資源は一定である。競合の結果、必要となる占有量が資源の量を超過した場合、ネットワーク内で、いずれかのパケットデータが消失することは明らかである。これを輻輳状態と呼ぶ。送信側では、パケット送出を規制することでこのデータの消失を防ぐ事が、前記1)、2)から必要となる。これを輻輳制御と呼ぶ。

【0013】一方、大容量の結果データを短時間に転送するためには、前記のような輻輳の制御を行なわないこともある。データ転送に先立ち、ネットワーク上の特定の伝送パスを事前に設定し、一定時間ネットワーク資源を占有する。伝送パスは、

①アプリケーション起動時に設定する場合（例えば、電話網などの回線交換）

②アプリケーション起動後、データ転送時に動的に伝送パスを設定する場合（例えば、RSVP: A New Resource ReSerVation Protocol）

の2通りがある。図3は、データ転送時に動的に伝送パスを設定するデータ転送装置の概略構成を示したものである。図中、輻輳回避型送信帯域規制手段106は、データ転送に先立ち、資源配分手段107に対して送信要求を行う。資源配分手段107は、送信要求を受信すると伝送パスを設定した後、データ転送の開始を指示する。伝送パスの設定に際し、伝送路の伝送路帯域が既知、かつ、占有できるものとして送出速度を決定する。伝送路の実際の使用状況は観測できない。

【0014】以上をまとめると、輻輳制御には次のふたつの方法があることになる。

【0015】①回復型輻輳制御（図2）：

前提：ネットワークの資源量と、他のデータ転送の資源使用状況とを知ることができない。

原理：アプリケーションの要求通りにネットワーク資源の使用を開始する。輻輳発生時のフィードバックによって使用量を低減する。輻輳発生後、輻輳状態から回復する。

特徴：各送信局は、任意のタイミングで遅滞なくデータ転送を開始できる。

特性：伝送効率は、ネットワークが低負荷の場合にはよいが、高負荷では低下する。伝送効率は、1回の転送データ量には依存しない。

従来例：ABR (Available Bit Rate)、TCP (Transmission Control Protocol)

②回避型輻輳制御（図3）：

前提：ネットワークの資源量と、他のデータ転送の資源使用状況とを知ることができる。

原理：アプリケーションの要求使用量の申告とネットワーク資源の未使用量から、各データ転送の資源使用量を配分する。輻輳状態を未然に回避する。

特徴：各送信局は、データ転送の要求発生からその開始までに遅延が生じる。

特性：伝送効率は、1回の転送データ量が大きいとよいが、小さいと低下する。伝送効率は、負荷に依存しない。

従来例：RSVP

近年、アプリケーションの要求するデータ量が大幅に増大している。しかし、LANという前提から、従来どおりのデータ量まで同時にカバーする必要がある。そこで、前記2種類の輻輳制御方式の利点を同一のネットワークで享受したいという要求がある。

【0016】例えば、

①従来のアプリケーションは、データ転送を回復型輻輳制御によって実行する

②大容量の処理結果データを広帯域で転送するアプリケーションは、データ転送を回避型輻輳制御によって実行する

というように、ひとつのネットワーク上で、アプリケーション、もしくは、要求転送データ量に応じて2種類の輻輳制御方式を使い分けることで、高い伝送効率を得ることができる。

【0017】このような状況においては、回復型輻輳制御に従うデータ転送（以下、輻輳回避型データ転送）と回避型輻輳制御に従うデータ転送（以下、輻輳回復型データ転送）とが、同一の伝送パスを流れる場合がある。

その場合、両者は伝送パス上の同一のネットワーク資源を使用する。使用されるネットワーク資源は、いずれか一方の輻輳制御方式によって単独で制御されるときと同様に、2種類の輻輳制御方式によって同時に制御されるときも、輻輳状態が制御される必要がある。

【0018】ところが、実際には、同一のネットワーク資源の輻輳を前記2種類の輻輳制御方式によって同時に制御する場合には、輻輳制御に破たんをきたし、伝送効率が著しく低下するという問題がある。

【0019】以下、クライアント-サーバ型データ転送装置を例に、この問題を説明する。

【0020】まず、クライアントーサーバ型データ転送装置に関して説明する。クライアントーサーバ型とは、図4にその概念を示すように、サービスの受益者（クライアント）110とサービスの提供者（サーバ）111の通信を表すモデルである。クライアントーサーバ型データ転送装置では、複数のクライアント110からのデータ転送が、ひとつのサーバ111に同時に集中することがある。

【0021】図5は、クライアントーサーバ型データ転送装置において、前記2種類の輻輳制御方式によるデータ転送を同時に収容する構成を示したものである。図中、クライアント110a上のアプリケーション（プログラム）102は輻輳回復型データ転送を行い、クライアント110bのアプリケーション102は輻輳回避型データ転送を行う。輻輳回避型データ転送と輻輳回復型データ転送は、サーバ111への入口である伝送路に集中し、混合される。サーバ111のデータ転送分離手段112は、前記2種類のデータ転送を受信して分離して、アプリケーションに渡す。

【0022】このとき、サーバ111への入口である伝送路において、輻輳回避型データ転送の伝送帯域と輻輳回復型データ転送の伝送帯域の総和が伝送路帯域を上回る過負荷状態が発生することがある。伝送路帯域を上回る分のデータはバッファ101aに一時的に滞留するが、バッファ101aは有限であるから、過負荷状態が継続するとバッファが溢れる。溢れたデータは消失する。これが輻輳である。

【0023】輻輳が発生すると、輻輳検知手段104が輻輳を検出し、輻輳回復型送信帯域規制手段105に通知することで、輻輳回復型送信帯域規制手段105がデータ転送の送出速度を低減する。

【0024】しかし、輻輳回避型送信帯域規制手段106は輻輳状態にかかわらず、データ転送開始時の送出速度のままデータ転送を継続する。

【0025】以上をまとめると、前記の問題が発生する理由は次のとおりである。

①回復型輻輳制御は、輻輳状態のフィードバックを受けて資源使用量を低減することで、資源を共有する。ところが、回避型輻輳制御は輻輳状態が発生しているにもかかわらず、あらかじめ配分された資源量を使用し続けようとする。これは回復型輻輳制御の原理に反し、回復型輻輳制御が破たんする。

②回避型輻輳制御は、ネットワーク資源の要求量の申告と資源の未使用量から使用量を配分して伝送パスを設定する。ところが、回復型輻輳制御は要求量の申告を行わずに、任意のタイミングでデータ転送を開始する。そのため、回避型輻輳制御は資源の未使用量を正しく把握することができない。これは回避型輻輳制御の原理に反し、回避型輻輳制御が破たんする。

【0026】このように、本来は共存しえない回復型輻

輳制御と回避型輻輳制御とによって、同一のネットワーク資源を使用してデータ転送を行うためには、次のふたつの方法がある。

【0027】第一は、ネットワーク資源を分離して、それぞれ異なる輻輳制御方式を適用する方法である。ネットワーク資源の分離は、

①物理的に分離する場合（例えば、無線伝送媒体による空間分割や周波数分割）

②論理的に分離する場合（例えば、ATM PVC-CBRのような伝送路帯域予約、ATM Forum TM Specification Version 4.0、ATM Forum ATM UNI Specification Version 3.1）の2通りがある。この場合、両者はひとつのネットワーク資源を共有しないで独立しているため、輻輳状態を制御できる。しかし、この方法は、一方の輻輳制御方式でのデータ転送が行われていない期間もネットワーク資源は確保されているため、ネットワークの使用効率が悪いという問題がある。

【0028】第二は、ネットワークを共有するすべてのデータ転送装置が、最大能力で同時にデータ転送を行っても、足りるだけのネットワーク資源をあらかじめ用意する方法である。例えば、伝送路を超広帯域化する。この場合、事実上輻輳は発生しない。しかし、この方法は、通常の使用では非常な低負荷であり、ネットワークの使用効率が悪いという問題がある。

【0029】どちらの方法も、ネットワークの使用効率が悪いという問題に加え、新たなアプリケーションの導入のたびにネットワークを再構築する必要がある。

【0030】

【発明が解決しようとする課題】本発明の目的は、ネットワークの分離や再構築を行うことなく、回復型輻輳制御と回避型輻輳制御とをコンピュータネットワーク上で共存させることである。ここで、共存とは、輻輳回避型データ転送と輻輳回復型データ転送とが、同一の伝送パスを経由し、かつ、伝送パス上の同一のネットワーク資源を共有するときも資源の輻輳状態を制御できることを意味する。

【0031】

【発明の基本原則】本発明に於ける、輻輳制御型データ転送装置は複数のクライアントとサーバの間でデータ転送をネットワークを介して行う。以下、本発明における共存を可能とする手法に関して説明する。

【0032】本装置に収容される前記回避型輻輳制御および回復型輻輳制御は、以下の手順をアプリケーションとの間で実行する。

・回避型輻輳制御は、アプリケーションのデータ転送要求時に、伝送路帯域から帯域使用量を順次割当てる。

・回復型輻輳制御は、アプリケーションのデータ転送要求時に、データ転送の帯域使用量を要求しない。そこ

で、輻輳回復型データ転送の帯域使用量の推定値に基づき、輻輳回避型データ転送の帯域使用量を割当てる。

【0033】輻輳回復型データ転送の帯域使用量は、伝送要求を受信する側のアプリケーション（以下、受信側アプリケーションと称する）の持つパラメータに基づいて推定される。

【0034】受信側アプリケーションが持つパラメータには、次の3つがある。

- ・アプリケーションの種別
- ・転送されるデータ量
- ・送信側アプリケーションが動作するコンピュータ種別。

【0035】コンピュータ種別からは、送信側アプリケーションの概略伝送帯域が推定される。伝送帯域は、帯域使用量の中心値 $W0$ とその時間変動量の2軸で推定される。まず、中心値 $W0$ の値は、帯域使用量の長時間で平均した値となり、例えば中心値 $W0$ は、以下のようにカテゴリーに分類される。

- ・低レート：中心値 $W0 = 1 \text{ Mbps}$ （リモートプロシージャコールなど）
- ・中レート：中心値 $W0 = 10 \text{ Mbps}$ （メール、ツールサービスなど）
- ・高レート：中心値 $W0 = 25 \text{ Mbps}$ （ファイル転送、文字原稿中心のプリントなど）

次に、その時間変動量を推定する。この場合、以下の条件のいずれかのケースでは、中心値のみで推定は完了する。それ以外のケースでは、中心値に加えて、時間変動量の推定を行う。このケースは、システム構成を決定する時点で選定される。

- ・送信局において送出速度の上限が規制（シェーピング）される。
- ・送信側アプリケーションが固定帯域でデータを送信する。

【0036】以上をまとめると、推定は、フェーズ1（中心値の推定）、フェーズ2（時間変動量の推定）の2段階に分割される。フェーズ1では、受信側アプリケーションの持つパラメータから、そのアプリケーションが行う輻輳回復型データ転送の中心値を推定する。尚、回復型データ転送が複数実行されている場合、推定総量は、中心値 $W0$ の累計 $\Sigma W0$ 、時間変動量を含めた総量 $\Sigma W(t)$ と定義される。

【0037】以上の手順を実行するために、本発明に係わる輻輳制御型データ転送装置は、次の5つの手段から構成される。図6は本発明の手段構成を示したものである。

【0038】①要求帯域取得手段1

要求帯域取得手段1は、データ転送に先だってアプリケーション間で交換される制御情報から、輻輳回避型データ転送の要求伝送帯域を取得する手段である。前記制御情報には、転送データ量、要求伝送時間などがある。

【0039】②伝送帯域推定手段2

伝送帯域推定手段2は、データ転送に先だってアプリケーション間で交換される制御情報から、輻輳回復型データ転送の伝送帯域を推定する手段である。前記制御情報には、転送データ量、送信側アプリケーションが動作するコンピュータの種別などがある。

【0040】図7は伝送帯域の推定について示したものである。伝送帯域の推定には次のふたつのフェーズがある。

- 10 フェーズ 1. 伝送帯域の中心値の選択
- フェーズ 2. 短時間変動の補正

③管理情報記憶手段3

管理情報記憶手段3は、現時点で伝送路帯域を使用中の輻輳回避型データ転送と、輻輳回復型データ転送との全てについて、その管理情報を記憶している手段である。前記管理情報には、各データ転送の伝送帯域が含まれる。

【0041】④伝送帯域割当て手段4

- 20 伝送帯域割当て手段4は、伝送路帯域の未使用量と、アプリケーションによって新規に要求されたデータ転送の要求伝送帯域とから、前記データ転送の実行の可否を判定すると共に、新規要求の輻輳回避型データ転送に対して伝送帯域を割当てる手段である。

【0042】図8は伝送帯域の割当て方法を模式的に示したものである。伝送路帯域の未使用量は、既知の伝送路帯域 WT から、管理情報記憶手段3に登録される輻輳回避型データ転送の伝送帯域の総計 WA と、管理情報記憶手段3に登録される輻輳回復型データ転送の総伝送帯域の推定量の総計 WB とを差し引くことで算出する。

- 30 【0043】⑤データ転送指示手段5

データ転送指示手段5は、新規のデータ転送要求について、アプリケーションに送信を指示する手段である。新規要求のデータ転送が輻輳回避型データ転送である場合、前記指示には伝送帯域割当て手段4で割当てられた伝送帯域が含まれる。

【0044】

【課題を解決するための手段】本発明によれば、上述の基本原則にしたがって、時分割交換ネットワークに接続された複数のコンピュータ上のアプリケーションが前記ネットワークを介してデータ転送を行い、特定のコンピュータに対して他の複数のコンピュータから単一の経路を介して同時にデータ転送が行われるネットワークシステムにおいて用いるデータ転送装置に、前記時分割交換ネットワークにおける輻輳を検出する輻輳検出手段と、前記特定のコンピュータによる前記アプリケーションのデータ転送要求の受信を監視して前記時分割交換ネットワークの資源使用総量を判別する資源使用量観測手段と、新規に要求されるデータ転送が第1の種類のデータ転送（輻輳回復型データ転送）のときに、前記時分割交換ネットワーク上で、前記データ転送の要求に応じてデ

ータ転送を開始し、前記輻輳検出のフィードバックにより、前記時分割交換ネットワークの資源使用量を制御する第1の転送制御手段と、新規に要求されるデータ転送が第2の種類のデータ転送（輻輳回避型のデータ転送）のときに、前記資源使用量観測手段で判別した資源使用量に基づき前記新規要求されるデータ転送に対する前記時分割ネットワークの資源使用可能量を求め、前記資源使用可能量の範囲で資源割当て量を通知する資源割り当て手段と、前記資源割当て量の通知に基づいて、前記新規に要求された第2の種類のデータ転送を前記資源割当て量の帯域で実行する第2の転送制御手段とを設けるようにしている。

【0045】この構成においては、データ転送要求を監視して、輻輳回復型のデータ転送に使用するネットワーク資源も決定するようにし、輻輳回復型のデータ転送と輻輳回避型のデータ転送とを同時に実行することができる。

【0046】また、この構成において、前記資源使用量観測手段は、前記第1のデータ転送の転送パスおよび前記第2のデータ転送の転送パスが集中する位置において、前記複数のコンピュータと前記特定のコンピュータ間の接続手順のための制御パケットを監視して前記時分割交換ネットワークの資源使用総量を判別するようにしてもよい。

【0047】また、前記資源使用量観測手段は、前記第1のデータ転送の要求を受信した時点でこのデータ転送に要するネットワーク帯域の中心値とその時間変動量を推定して、前記第1のデータ転送による前記時分割交換ネットワークの資源使用量を集計し、さらに前記第2のデータ転送による前記時分割交換ネットワークの資源使用量を集計し、これら集計した値を加算して前記時分割交換ネットワークの資源使用総量を判別するようにしてもよい。

【0048】また、前記第1のデータ転送を、対応するアプリケーションの種別および対応するアプリケーションの処理パラメータの少なくとも1つから予め設定した複数のカテゴリーに分類し、前記カテゴリーに対応して予め設定した伝送帯域値を前記第1のデータ転送のネットワーク帯域の中心値と推定するようにしてもよい。

【0049】また、前記第1のデータ転送に要するネットワーク帯域の時間変動量は、前記前記資源使用量観測手段において集計したデータ転送の総量の短時間変動量観測値に、一定の係数を乗じることにより決定するようにしてもよい。

【0050】また、前記資源割り当て手段は、前記第2のデータ転送の許可時間間隔を一定の時間に分割し、この時間間隔ごとに資源割当てを実行し、前記資源使用量観測手段は一定のデータパケット到着時間間隔に基づいて前記データ転送の総量の短時間変動量観測値を算出するようにしてもよい。

【0051】また、前記他の複数のコンピュータはクラ

アントであって、前記特定のコンピュータはサーバーであってもよい。また、前記特定のコンピュータはプリントサーバーであってもよい。また、前記特定のコンピュータは前記ネットワーク内のスイッチであってもよい。また、前記特定のコンピュータは、前記データ転送の相手先を選択するネットワーク通信制御サーバであってもよい。

【0052】また、本発明によれば、ネットワークにより接続された複数のコンピュータの間のデータ転送を行うデータ転送装置に、前記ネットワークにおける輻輳を検出する輻輳検出手段と、前記輻輳検出手段による輻輳検出をフィードバックして前記ネットワークの資源利用量を制御しながら前記コンピュータ間のデータ転送を行う第1の転送制御手段と、前記データ転送が集中する経路における前記ネットワークの資源使用量を観測する資源使用量観測手段と、前記資源使用量観測手段で観測された資源使用量に基づいて、使用可能な範囲で前記ネットワークの資源を割り当てて、前記コンピュータ間のデータ転送を実行する第2の転送制御手段とを設けるようにしている。

【0053】この構成においても、データ転送が集中する経路においてネットワークの資源利用量を観測し、これにより輻輳回復型のデータ転送の資源利用量をも決定し、これにより、輻輳回復型のデータ転送と輻輳回避型のデータ転送とを同時に実行させることができる。

【0054】また、本発明は、データ転送方法としても実現可能であり、またコンピュータ・プログラム製品としても実装可能である。

【0055】

【発明の実施の形態】本発明の実施の形態を図9に示す。図9において、図5または図6と対応する個所には対応する符号を付した。ここでは、まず、本発明のデータ転送手法が実装される位置に関して説明する。

【0056】本発明に係わる、輻輳制御型データ転送装置は、輻輳回復型データ転送と輻輳回避型データ転送が同時に集中して輻輳が発生しうるコンピュータ（例えばサーバ）のアプリケーションと、データ転送制御プログラムの間に位置する。即ち、ネットワーク上で、実質上データ転送要求が集中する場所である。そこで、複数のデータ転送が集中する位置に着目する。

【0057】クライアントーサーバ型のアプリケーションでは、多数のクライアントとのデータ転送がサーバに集約される。この場合、スイッチ型のネットワークでは、輻輳回復型データ転送と輻輳回避型データ転送とは、サーバとネットワークとを接続する伝送路に集中する。このとき、伝送路の一方の端であるネットワークの接続ポート、もしくは、もう一方の端であるサーバの入口において、伝送路に集中した全てのデータ転送を観測できる。

【0058】このように、データ転送が集中して輻輳が

発生する位置が局在し、かつ、あらかじめその位置を設計できるため、その位置において本発明を実施できる。

【0059】図9において、クライアント110aからサーバ111に対して輻輳回復型のデータ転送が行われると、伝送帯域推定手段2がこれを監視して、資源利用量を推定する。推定した資源利用量は間情報記憶手段3に記憶される。他方、クライアント110bからサーバ111に対して輻輳回避型のデータ転送を行うには、輻輳回避型送信帯域規制手段106が輻輳回避型のデータ転送要求を送出する。要求帯域取得手段1はこのデータ転送要求から要求帯域を判別する。伝送帯域割当て手段4は、管理情報記憶手段3にアクセスし、輻輳回復型データ転送に使用されている資源推定値および、他の輻輳回避型データ転送に用いられている資源を把握し、これに応じて、先の輻輳回避型データ転送要求に対して帯域を割り当てる。データ転送指示手段5は資源割り当てに応じて輻輳回避型送信帯域規制手段106に送信指示を送る。

【0060】このようにして、輻輳回復型のデータ転送と輻輳回避型のデータ転送とを同時に行うことが可能になる。

【0061】

【実施例】

【実施例1】次に、クライアントーサーバ型アプリケーションのデータ転送装置を例にとって、本発明の一実施例を説明する。

【0062】図10は実施例1のデータ転送装置の全体構成を示したものである。図中、データ転送はクライアント110a、110bからサーバ111に対して行われ、クライアント110a、110bが送信局、サーバ111が受信局である。データ転送装置は、輻輳回復型送信帯域規制手段105を有するクライアント110aと、輻輳回避型送信帯域規制手段106を有するクライアント110bと、データ転送分離手段112と、データ転送観測部11と、資源使用量推定部12と、資源配分部13とを有するサーバ111と、それらを相互に接続してパケットによるデータ転送を実行する、輻輳検知手段104と、バッファ101aとを有するパケット交換スイッチ101bから成る。

【0063】サーバ111のアプリケーション102は、輻輳回復型データ転送を用いたデータ転送を受け付けたことを契機に、データ転送観測部11のデータ転送検知機能14（図11）に対して、受信開始信号を通知する。また、同データ転送が完了したことを契機に、データ転送観測部11のデータ転送検知機能14に対して、受信完了信号を通知する。

【0064】輻輳回復型送信帯域規制手段105、輻輳検知手段104は、従来技術のABRと同様である。一方、輻輳回避型送信帯域規制手段106は、本出願人が先に提案した競合制御機能を備えたデータ伝送手法（特

願平8-17949号）と同等である。すなわち、特願平8-17949号の手法においては、各データ伝送の開始時刻、終了時刻、送信端末、受信端末、経路、使用帯域の変化等の伝送状況を確定的に掌握し、これに基づいて、輻輳を回避するように一元的にデータ転送を制御している。ただし、この実施例では、従来の輻輳回復型データ転送も含まれ、この種類のデータ伝送についてはその状況をすべて確定的に掌握することができない。したがって、輻輳回復型データ転送については資源使用量を推定し、この推定値を利用して、輻輳回避型データ転送にネットワーク資源の割り当てを行う。輻輳回避型送信帯域規制手段106はアプリケーション102によってデータ転送要求が発生すると、資源配分部13に対して、送信要求を通知する。

【0065】資源配分部13から、送信指示が通知されると、その指示内容の送信帯域でデータ転送を実行する。データ転送が完了すると、資源配分部13に対して、送信完了を通知する。

【0066】以下、各部の構成、動作に関して説明する。

【0067】図11は実施例1のデータ転送観測部11の構成を、図12は動作を示したものである。

【0068】データ転送観測部11には、データ転送検知機能14と、データ転送計数機能15と、伝送帯域統計機能16がある。

【0069】適応業務対応のプログラム102は、輻輳回復型データ転送を用いたデータ転送を受け付けたことを契機に、データ転送観測部11のデータ転送検知機能14に対して、受信開始信号を通知する。また、同データ転送が完了したことを契機に、データ転送観測部11のデータ転送検知機能14に対して、受信完了信号を通知する。

【0070】データ転送検知機能14は、初期化ののち（S11）、適応業務対応のプログラム102から受信開始信号を通知されると、データ転送計数機能15に対して転送開始信号を通知するとともに、データ転送管理テーブルに登録する（S12、S13）。図13は実施例1のデータ転送管理テーブルの形式を示したものである。つまり、データ転送管理テーブルには現時点で継続中のデータ転送がエントリされている。

【0071】また、データ転送検知機能14は、アプリケーション102から受信完了信号を通知されると、データ転送管理テーブルのエントリを検査する。エントリに一致する場合、データ転送計数機能15に対して転送完了信号を通知するとともに、データ転送管理テーブルのエントリを削除する（S12、S15）。

【0072】データ転送計数機能15は、計数カウンタを有し、データ転送検知機能14から転送開始信号を通知されたことを契機にカウンタ値を増加し（S14）、転送完了信号を通知されたことを契機にカウンタ値を減

少させる(S16)。すなわち、計数カウンタは生存データ転送数を保持している。データ転送計数機能15は、カウンタ値が変化する毎に、伝送帯域統計機能16および資源使用量推定部12に対して生存データ転送数を通知する(S17)。

【0073】伝送帯域統計機能16は、伝送帯域統計テーブルと、計数カウンタと、タイマとを有している。

【0074】伝送帯域統計テーブルは、生存データ転送数毎に累計受信時間と累計受信パケット数を保持する。図14はその形式を示したものである。

【0075】伝送帯域統計機能16は、データ転送計数機能15から生存データ転送数が通知されることを契機にタイマをリセットして、再スタートする。すなわち、タイマは、生存データ転送数が一定である期間の時間長を計測している。リセットに際し、伝送帯域統計テーブルに生存データ転送数毎に保持されている累計受信時間にタイマ値を加算する。

$$\text{平均伝送帯域 [bps]} = \frac{\text{累計受信パケット数} \times \text{固定パケット長 [bit]}}{\text{累計受信時間 [sec]}} \quad (1)$$

データ転送ひとつ当たりの伝送帯域の長時間平均値は、ユーザ・インタフェースを介して、ユーザがあらかじめ入力した固定値を用いるように構成してもよい。あるいは、ネットワーク管理情報を用いてもよい。

【0079】さらに別の実施例では、データ転送ひとつ当たりの伝送帯域の長時間平均値の代わりに、データ転送速度の上限を規制するシェーピング速度を用いてもよい。

【0080】図15は実施例1の資源使用量推定部12の構成を、図16は動作を示したものである。

$$R_{used} = N \times R_{ave}$$

ここで、

R_{used} : 輻輳回復型データ転送の使用帯域の推定値 [bps]
 R_{ave} : 各輻輳回復型データ転送の使用帯域の長時間平均値 [bps]
 N : データ転送生存数

短時間変動補正機能18は、総使用帯域算出機能17から輻輳回復型データ転送の総使用帯域の推定値 R_{used} が通知されると、それに対して短時間変動分の補正を行う(S24)。

【0084】補正方法を説明する。データ転送の、ある計測期間中の伝送帯域の時間平均値は、計測期間が長ければ長いほど長時間平均値(統計平均値)に近づく。逆

$$\text{制御周期時間 } T [\text{sec}] = \frac{\text{要求転送データ量 [bit]}}{\text{要求送信帯域 [bps]}} \quad (3)$$

そこで、制御周期時間長に反比例する短時間変動分を、伝送帯域の増加方向にマージンとしてとる。補正後の輻

【0076】伝送帯域統計機能16は、輻輳回復型データ転送のパケットを受信したことを契機として、計数カウンタのカウンタ値を増加させる。すなわち、計数カウンタは受信パケット数を計数している。伝送帯域統計機能16は、タイマがリセットされたことを契機として、伝送帯域統計テーブルに生存データ転送数毎に保持されている累計受信パケット数にカウンタ値を加算し、計数カウンタをリセットする。

【0077】伝送帯域統計機能16は、データ転送計数機能15から生存データ転送数が通知されることを契機に、伝送帯域統計テーブルに生存データ転送数毎に保持されている累計受信時間と累計受信パケット数から、生存データ転送数に対応する平均伝送帯域 [bps] を次式によって算出し、資源使用量推定部12に対して通知する。

【0078】
 【数1】

【0081】資源使用量推定部12には、総使用帯域算出機能17と、短時間変動補正機能18がある。

【0082】総使用帯域算出機能17は、初期化の後(S21)、データ転送観部11から、生存データ転送数 N とデータ転送ひとつ当たりの伝送帯域の長時間平均値 R_{ave} が通知されると、輻輳回復型データ転送の総使用帯域 R_{used} を(2)式により算出する(S22、S23)。

【0083】
 【数2】

に、計測期間が短い場合には短時間変動分が誤差として無視できない。輻輳回避制御の制御周期時間内の輻輳回復型データ転送の使用帯域を推定することが目的であるから、(3)式で示される輻輳回避制御の制御周期時間が、前記計測期間に相当する。

【0085】
 【数3】

輻輳回復型データ転送の使用帯域の推定値 R_{used}' の算出式を(4)式に示す。

【0086】

【数4】

$$R_{used}' = R_{used} \times \left(1 + \frac{a}{T}\right) \quad (4)$$

ここで、

R_{used}' : 輻輳回復型データ転送の使用帯域の推定値 [bps]
(補正後)
 R_{used} : 輻輳回復型データ転送の使用帯域の推定値 [bps]
(補正前)
 a : 補正係数
 T : 輻輳回避制御の制御周期時間 [sec]

短時間変動補正機能18は、補正を行って輻輳回復型データ転送の使用帯域 R_{used}' を算出すると、それを使用帯域記憶メモリに記憶する(S25)。使用帯域記憶メモリを更新する度に、資源配分部13に資源配分信号を通知する(S26)。

【0087】他の実施例では、送信局毎に記憶媒体の入出力速度が異なるという前提の下で、送信局毎に異なるデータ転送ひとつ当たりの伝送帯域の長時間平均値 R_{ave} 20

$ve(i)$ ($i=0, 1, 2, \dots, m$; i は送信局を識別する識別子)から輻輳回復型データ転送の使用帯域を推定してもよい。その場合、輻輳回復型データ転送の使用帯域の推定値 R_{used} の算出式は(5)式となる。

【0088】

【数5】

$$R_{used} = \left(1 + \frac{a}{T}\right) \sum_{i=1}^N R_{ave}(i) \quad (5)$$

ここで、

R_{used} : 輻輳回復型データ転送の使用帯域の推定値 [bps]
 $R_{ave}(i)$: 各輻輳回復型データ転送の使用帯域の長時間平均値
(送信局毎に固定) [bps]
 N : データ転送生存数
 a : 補正係数
 T : 輻輳回避制御の制御周期時間 [sec]

図17は実施例1の資源配分部13の構成を、図18、図19、図20は動作を示したものである。

【0089】資源配分部13には、送信待ち行列機能19、送信帯域決定機能20、送信帯域指示機能21がある。

【0090】資源配分部13には、自らが各輻輳回避型データ転送に対して既に通知した送信帯域値を記憶する資源使用量管理テーブルを有している。図21は資源使用量管理テーブルの形式を示したものである。

【0091】送信待ち行列機能19は、FIFO (First In, First Out) 待ち行列を有している。送信局から送信要求を通知されると、その内容を待ち行列にキューイングする(S31、S32、S33)。送信待ち行列機能19は、待ち行列内の待ち要求の有無を送信帯域決定機能20に対して常時表示する。

【0092】また、送信局から送信完了を通知されると、資源使用量管理テーブル中の対応するエントリを削除し、送信帯域決定機能20に対してデータ転送完了信

号を通知する(S31、S34、S35、S33)。

【0093】送信帯域決定機能20は、伝送路帯域値を固定値として保持している。

【0094】送信帯域決定機能20は、資源使用量推定部12から輻輳回復型データ転送の使用帯域が通知されると、輻輳回避型データ転送への送信帯域の割当て、ならびに送信局から送信完了を通知されることを契機として、エントリを更新される。

40 【0095】送信帯域決定機能20は、伝送路帯域値 R_{link} から、輻輳回復型データ転送の使用帯域の推定値 R_{used} と、実行中の輻輳回避型データ転送に対して既に配分した帯域値 R_i ($i=0, 1, 2, \dots, m$; m は実行中の輻輳回避型データ転送数)の総計とを差し引くことで、空き帯域値 R_{max} を算出する。

【0096】算出式を(6)式に示す。

【0097】

【数6】

$$R_{\max} = R_{\text{link}} - R_{\text{used}} - \sum_{i=1}^m R_i \quad (6)$$

ここで、

- R_{\max} : 空き帯域の算出値
 R_{link} : 伝送路帯域値(固定) [bps]
 R_{used} : 輻輳回復型データ転送の使用帯域の推定値 [bps]
 R_i : 実行中の各輻輳回避型データ転送への配分帯域 [bps]
 m : 実行中の輻輳回避型データ転送数

送信待ち行列機能19においてキューイングされている先頭の送信要求に対して、算出された空き帯域値の範囲内で、要求送信帯域に基づいて送信帯域を割り当てる。すなわち、送信待ち行列の先頭の送信要求を取り出してこれを解析し要求送信帯域を求める(S41、S42)。そして、空き帯域値を算出し、割り当て可能であれば、送信帯域値を決定し(S43、S44)、資源使用量管理テーブルにエントリを追加し(S45)、送信帯域指示機能21に送信帯域値を通知する(S46)。空き帯域が少ないせいで所望の帯域を割り当てられない場合は、他のデータ転送が完了して割り当てられるようになるまで、送信待ち行列機能19に保持しつづける(S47)。

【0098】送信帯域指示機能21は、送信帯域決定機能20において決定した送信帯域値を含む送信指示を生成し、送信要求の送信元である送信局に対して通知する(S51~S53)。

【0099】本実施例において、輻輳回避型データ転送と輻輳回復型データ転送とがクライアント110からサーバ111に対して同時に実行された場合に、サーバ111の入口である伝送路において輻輳回避型データ転送の伝送効率の評価を行った。図22は評価結果の一例を示したものである。図中、代表的なサービス例えばプリントサービス等を想定し、以下に設定している。

①回避型データ転送数(即ち、同時伝送データ本数) : 1本。最大伝送路帯域までの帯域を使用。

②回復型データ転送数 : ネットワークの空き帯域に依存して、最低1本から、最大20本まで。更に、伝送帯域は、中心値を伝送路帯域の4%。各々のデータ転送帯域の時間変動は、ポアソン過程に従うものとした。

③輻輳の判断基準 : データ消失率 10^{-7} 以下
 先ず、従来の構成のままの場合、帯域乗利用効率は15%に止まる。資源使用量推定部において、フェーズ1(平均値)のみで推定した場合、70%の空き帯域利用効率を得た。

【0100】更に、フェーズ2(時間変動量補正)まで含めると95%の空き帯域利用効率を得た。この時、第4式における制御間隔Tを一定の回避型輻輳制御の伝送間隔とし、補正係数aを実験的に最適化した。また、複数の輻輳回復型データ転送トラヒックの集合の時間変

動分布は、実験的に一定分布とした場合が、最適、即ち固定レートとした場合が最適であった。

【0101】なお、上述の実験からも容易に理解できるように、輻輳回避型データ転送の制御期間を、輻輳回避型データ転送に許容された期間を分割した一定の時間間隔とし、この時間間隔ごとに割り当てを行うようにしてもよい。また、一定のバケットの到着時間間隔に基づいて短時間変動量観測値を算出するようにしてもよい。

【0102】また、図10において、111aは、この実施例のデータ転送制御をサーバ111に実行させるためにインストールされるプログラムを保持しているプログラム製品であり、これを用いて本実施例の手法をサーバ111に実装できる。

【0103】[実施例2] 実施例2では、クライアント-サーバ型アプリケーションのデータ転送装置を例にして説明する。図23は実施例2の全体構成を示す。データ転送はクライアント110a、110bからサーバ111に対して行われ、クライアント110a、110bが送信局、サーバ111が受信局である。

【0104】実施例2の観測点に関して説明する。実施例2はクライアント-サーバ型アプリケーションであって、輻輳回復型データ転送と輻輳回避型データ転送とが集中して輻輳が発生しうる位置は、受信局への入口であることから、中継スイッチ(パケット交換スイッチ101b)の接続ポートに観測点を設ける。

【0105】また、中継スイッチの配置を設計することによって、輻輳回復型データ転送と輻輳回避型データ転送とが集中して輻輳が発生しうる位置を、この一点だけとするように設計している。

【0106】実施例2の観測対象に関して説明する。

【0107】アプリケーションプログラムの処理結果の伝送は、通常、データ転送制御ソフトウェアにおいて接続指向のデータ転送により実行される。ここで接続指向のデータ転送とは、データ転送毎に接続の確立、維持、解放と呼ばれる制御手順を制御バケットによって行うデータ転送をいう。

【0108】すなわち、データ転送制御ソフトウェアの接続の確立と解放の間に、アプリケーションプログラムの情報交換、および、アプリケーションプログラムの処理結果の伝送が実行される。

【0109】そこで、データ転送制御ソフトウェアの接続の確立と解放のためのパケットを観測することは、アプリケーションプログラムの情報交換を観測することの代用となる。

【0110】図24は、従来のTCPを用いた接続動作を、UNIXオペレーティングシステム上の一般的プリントアプリケーションであるlpr（プリント）を例に示したものである。図中、送信側クライアントアプリケーションと受信側サーバアプリケーションとの間で、ファイルデータの転送に先だって、TCPを用いた接続が10
行われる。接続確立および接続解放の制御パケットも前記観測点に集中する。

【0111】そこで、輻輳回復型データ転送の接続確立および接続解放の制御パケットを検出することで、データ転送が資源を使用している継続期間を検知することができる。またデータ転送が資源を使用している継続期間を検出することもできる。

【0112】つぎに図23に戻り、データ転送装置の全体構成を説明する。図中、実施例2のデータ転送装置は、輻輳回復型送信帯域規制手段105を有するクライアント110aと、輻輳回避型送信帯域規制手段106を有するクライアント110bと、データ転送分離手段112とを有するサーバ111と、それらを相互に接続してパケットによるデータ転送を実行する、バッファ101aと、輻輳検知手段104と、データ転送抽出部22と、データ転送観測部11と、資源使用量推定部12と、資源配分部13とを有するパケット交換スイッチ101bとからなる。

【0113】実施例2の輻輳回復型送信帯域規制手段105、輻輳検知手段104、輻輳回避型送信帯域規制手段106、データ転送観測部11と、資源使用量推定部12および資源配分部13は実施例1と同等であるので説明を省略する。

【0114】図25は実施例2のデータ転送抽出部22の構成を、図26は動作を示したものである。データ転送抽出部22には、パケット分離機能23、接続検知機能24がある。

【0115】一般に、データ転送制御ソフトウェアは階層化されたデータ転送プロトコルとして実装されている。階層化されたデータ転送プロトコルによって生成されたパケットは、プロトコルの制御情報を保持しているヘッダと呼ばれる構造を有している。図27に示すヘッダ構造を示す。すなわち、階層(N)のプロトコルの情報には、1つ上位の階層(N+1)プロトコルを指定する情報が含まれており、これがプロトコル識別子である。図28は従来のTCP/IP(Internet Protocol)によるコネクション、および、プロトコル指定を示したものである。

【0116】データ転送制御ソフトウェアにおいて、2種類の輻輳制御方式は異なるプロトコルとして実装され50

る。そこで、プロトコル識別子を検査することにより、データ転送を2種類の輻輳制御方式に分別できる。

【0117】また、アプリケーション間の接続の確立と解放のための情報交換は、ヘッダの特定の部位(フィールド)の値によって行われる。そこで、ヘッダの特定のフィールドを検査することにより、接続の確立と解放の手順を観測できる。

【0118】データ転送抽出部22は、図29に示す形式のプロトコル情報テーブルを保持している。

【0119】パケット分離機能23は、伝送路からパケットを受信すると、受信パケットのプロトコル識別子を前記プロトコル情報テーブル(図29)に照らし合わせる(S61、S62)。プロトコル識別子がプロトコル情報テーブル中の輻輳回復型データ転送かつ接続指向データ転送に対応するエントリに一致する場合に、受信パケットを接続検知機能24に渡す(S63)。それ以外の場合は受信パケットを破棄する(S67)。

【0120】接続検知機能24は、パケット分離機能23から輻輳回復型データ転送かつ接続指向データ転送の受信パケットを受信すると、前記プロトコル識別子テーブルに照らし合わせる。受信パケットが接続指向の輻輳回復型データ転送のプロトコル識別子に対応する接続確立手順を示すパケット内容である場合、受信開始信号を資源使用量推定部12に通知するとともにデータ転送管理テーブルに登録する(S68、S69)。図30にデータ転送管理テーブルの形式を示す。つまり、データ転送管理テーブルには現時点で接続中のデータ転送がエントリされている。他方、受信パケットが接続指向の輻輳回復型データ転送のプロトコル識別子に対応する接続解放手順を示すパケット内容である場合、データ転送管理テーブルのエントリを検査する(S65)。エントリに一致する場合、受信完了信号を資源使用量推定部に通知するとともに、データ転送管理テーブルのエントリを削除する(S65、S66)。

【0121】この実施例においても実施例1と同様に輻輳回復型データ転送と輻輳回避型データ転送を同時に行うことが可能になる。

【0122】

【発明の効果】本発明によれば、回復型輻輳制御と回避型輻輳制御とがネットワーク上に同時に存在して、ひとつの伝送路を動的に共有し、かつ、負荷が大きい時も、データ消失なく、高い伝送効率をえられる効果がある。これにより、ネットワークの分離や再構築をすることなく、従来のアプリケーションは回復型輻輳制御、大容量かつ高負荷のデータ転送を行うアプリケーションは回避型輻輳制御というように、アプリケーションに応じて使い分けることが可能となる。

【図面の簡単な説明】

【図1】 コンピュータネットワークの概略構成図である。

- 【図2】 従来のデータ転送装置の概略構成図である。
- 【図3】 輻輳の制御を行わないデータ転送装置の概略構成図である。
- 【図4】 クライアント-サーバモデルの概念図である。
- 【図5】 クライアント-サーバ型データ転送装置の構成説明図である。
- 【図6】 本発明の手段構成説明図である。
- 【図7】 本発明の伝送帯域推定手段の説明図である。
- 【図8】 本発明の伝送帯域割当て方法の概略説明図である。 10
- 【図9】 本発明の実施の形態説明図である。
- 【図10】 実施例1の全体構成説明図である。
- 【図11】 実施例1のデータ転送観測部の構成説明図である。
- 【図12】 実施例1のデータ転送観測部の動作説明図である。
- 【図13】 実施例1のデータ転送管理テーブルの説明図である。
- 【図14】 実施例1の伝送帯域統計テーブルの説明図 20
- 【図15】 実施例1の資源使用量推定部の構成説明図である。
- 【図16】 実施例1の資源使用量推定部の動作説明図である。
- 【図17】 実施例1の資源配分部の構成説明図である。
- 【図18】 実施例1の資源配分部の送信待ち行列機能の動作説明図である。
- 【図19】 実施例1の資源配分部の送信帯域決定機能 30
- 【図20】 実施例1の資源配分部の送信帯域指示機能の動作説明図である。
- 【図21】 実施例1の資源使用量管理テーブルの説明図である。
- 【図22】 実施例1の評価説明図である。
- 【図23】 実施例2の全体構成説明図である。
- 【図24】 従来技術TCPを用いた接続動作の例（1 pr）である。
- 【図25】 実施例2のデータ転送観測部の構成説明図 40
- 【図26】 実施例2のデータ転送観測部の動作説明図である。

【図27】 階層型データ転送プロトコルによるバケット構造説明図である。

【図28】 従来技術TCP/IPによるコネクション、および、プロトコル指定の説明図である。

【図29】 実施例2のプロトコル情報テーブルの説明図である。

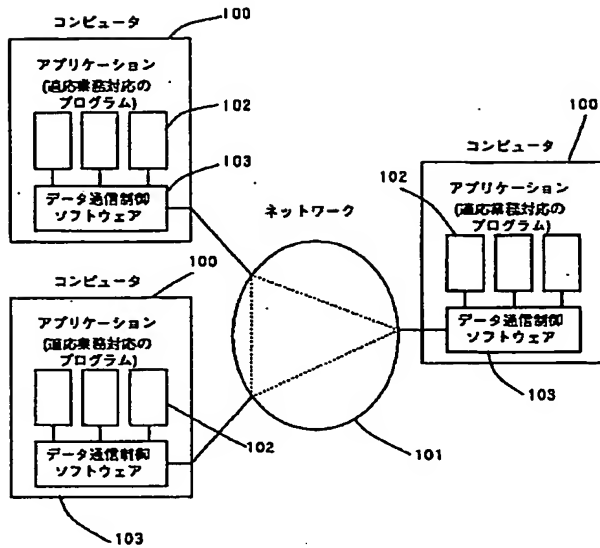
【図30】 実施例2のデータ転送管理テーブルの説明図である。

【符号の説明】

- 1 要求帯域取得手段
- 2 伝送帯域推定手段
- 3 管理情報記憶手段
- 4 伝送帯域割当て手段
- 5 データ転送指示手段
- 11 データ転送観測部
- 12 資源使用量指定部
- 13 資源配分部
- 14 データ転送検知機能
- 15 データ転送計数機能
- 16 伝送帯域統計機能
- 17 総使用帯域算出機能
- 18 短時間変動補正機能
- 19 送信待ち行列機能
- 20 送信帯域決定機能
- 21 送信帯域指示機能
- 22 データ転送抽出部
- 23 バケット分離機能
- 24 接続検知機能
- 100 コンピュータ
- 101 ネットワーク
- 101a バッファ
- 101b バケット交換スイッチ
- 102 アプリケーション（適用業務対応のプログラム）
- 103 データ通信制御ソフトウェア
- 104 輻輳検知手段
- 105 輻輳回復型送信帯域規制手段
- 106 輻輳回避型送信帯域規制手段
- 107 資源配分手段
- 110a、110b クライアント
- 111 サーバ
- 112 データ転送分離手段

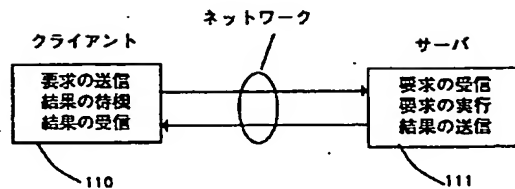
【図1】

コンピュータ・ネットワークの概略構成図



【図4】

クライアント-サーバモデルの概念図



【図13】

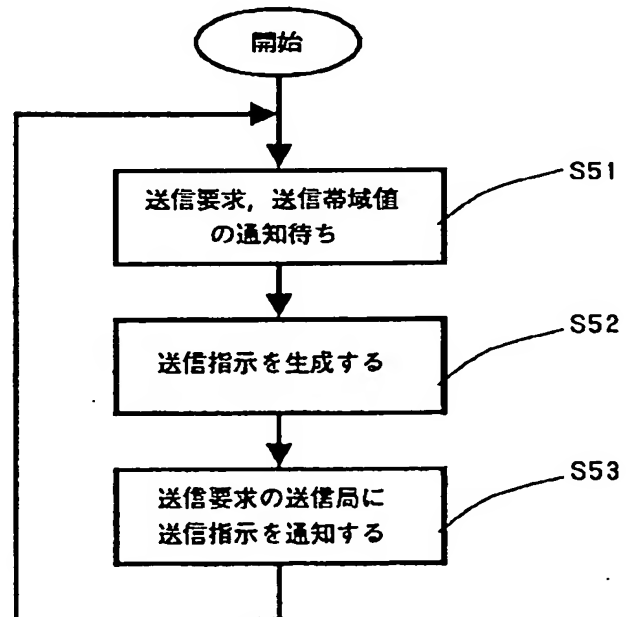
実施例1のデータ転送管理テーブルの説明図

	アプリケーション名	送信局アドレス
1		
2		
3		

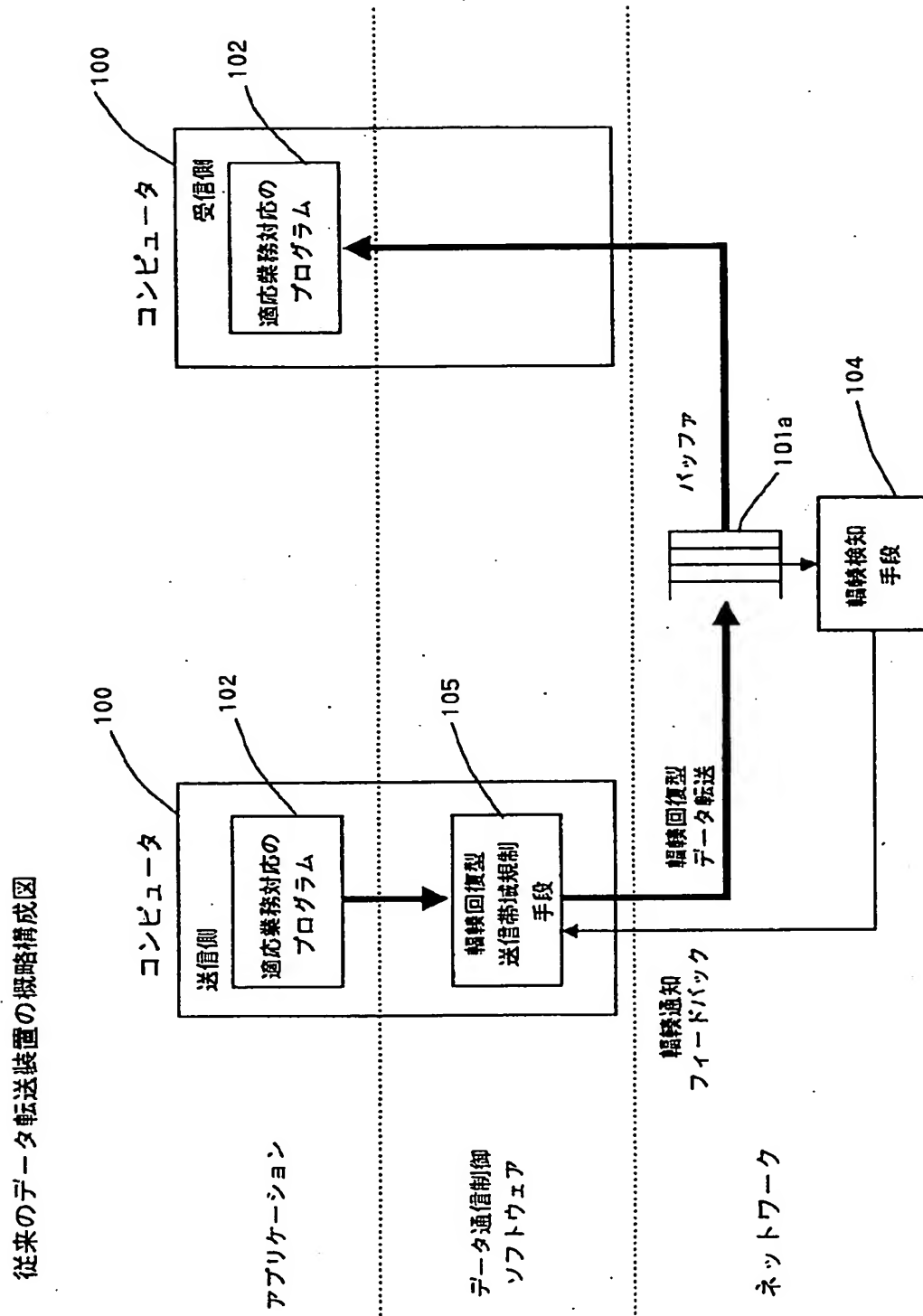
⋮

【図20】

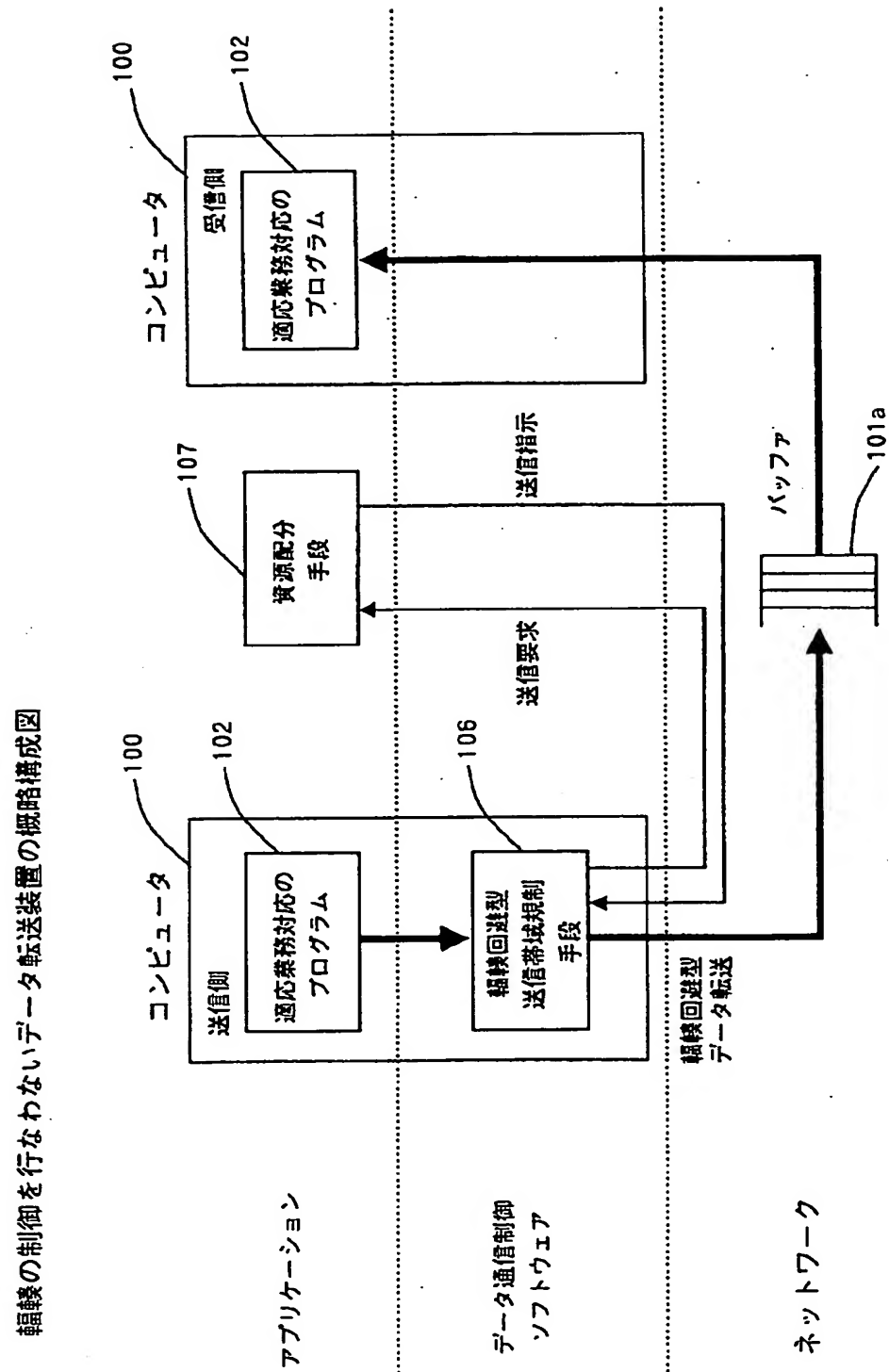
実施例1の資源配分部の送信帯域指示機能の動作説明図



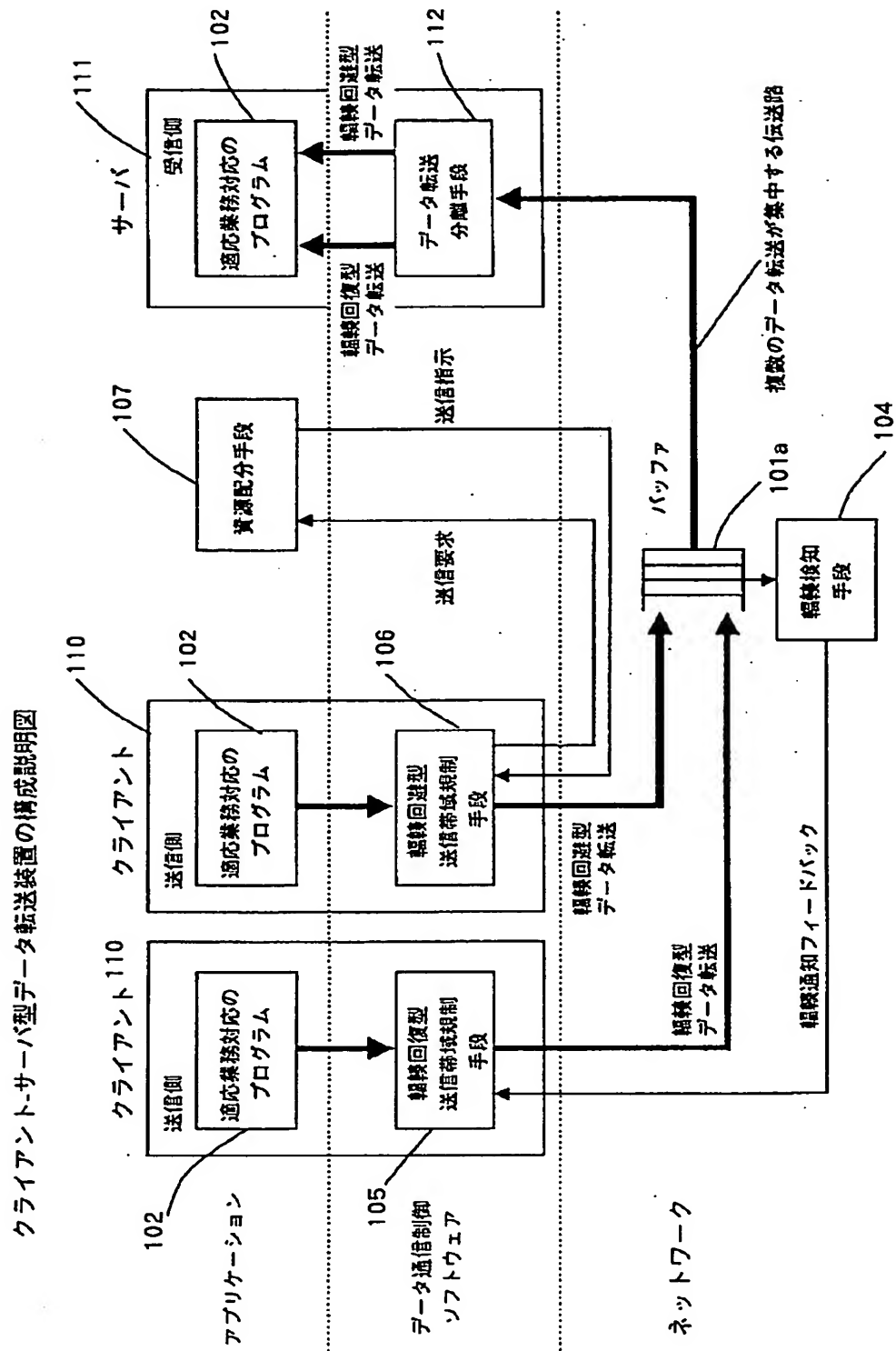
【図2】



【図3】

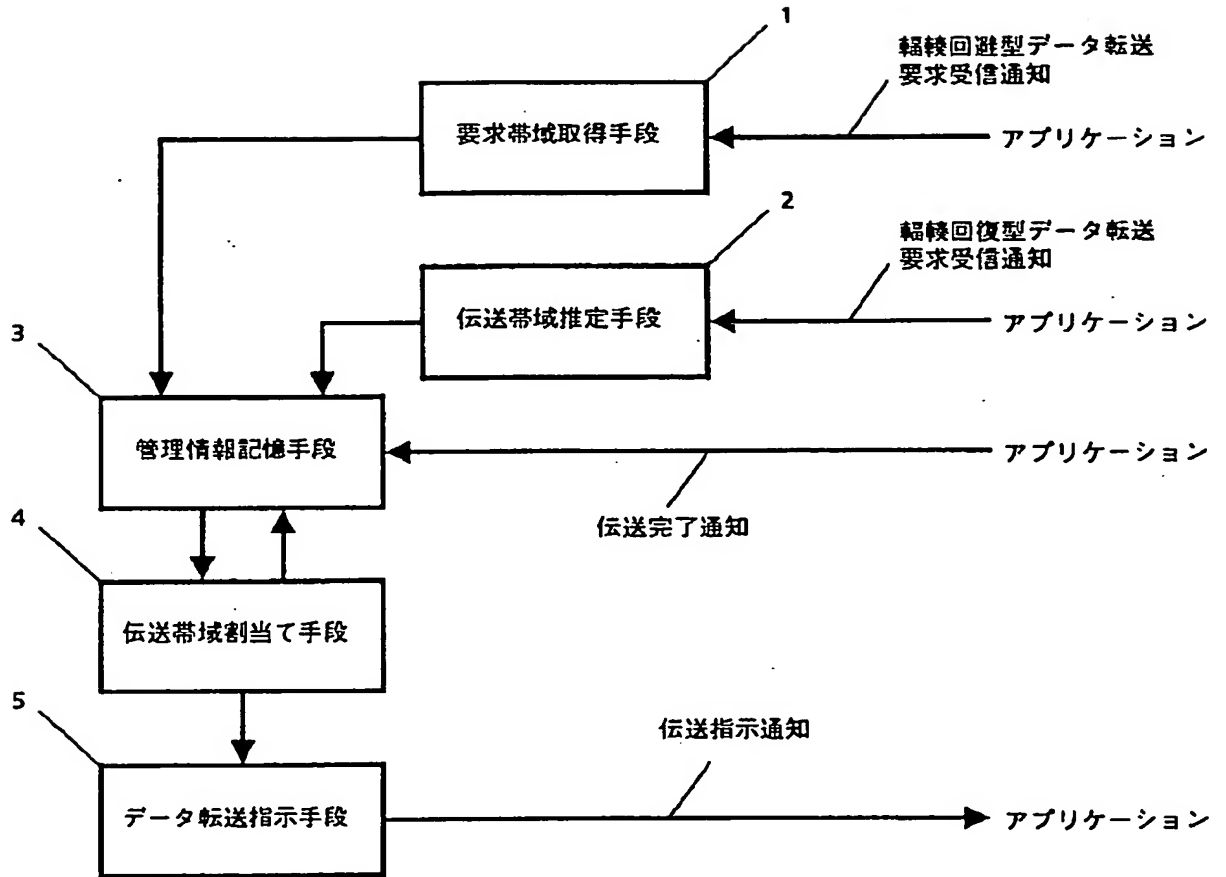


【図5】



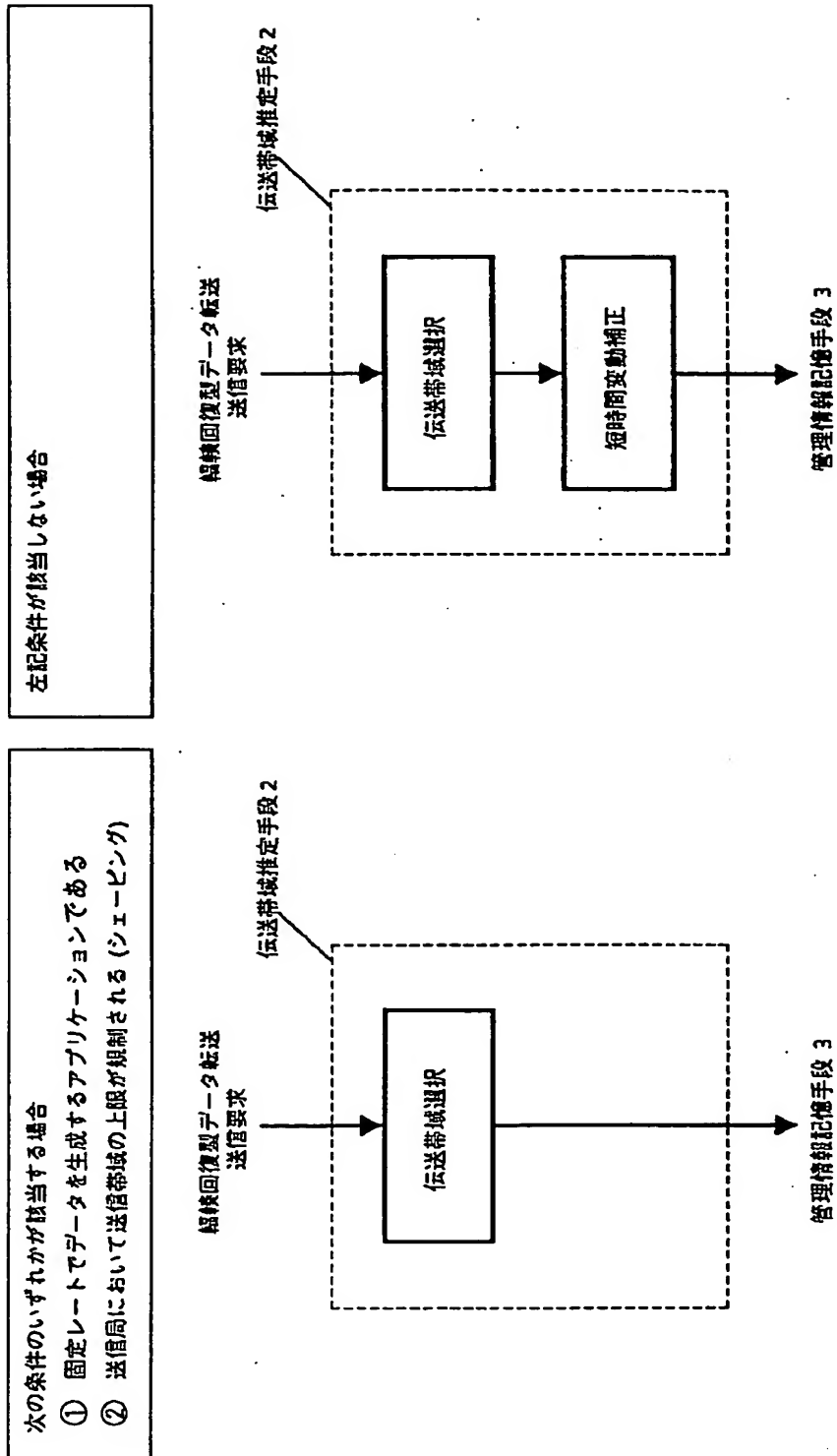
【図6】

本発明の手段構成説明図

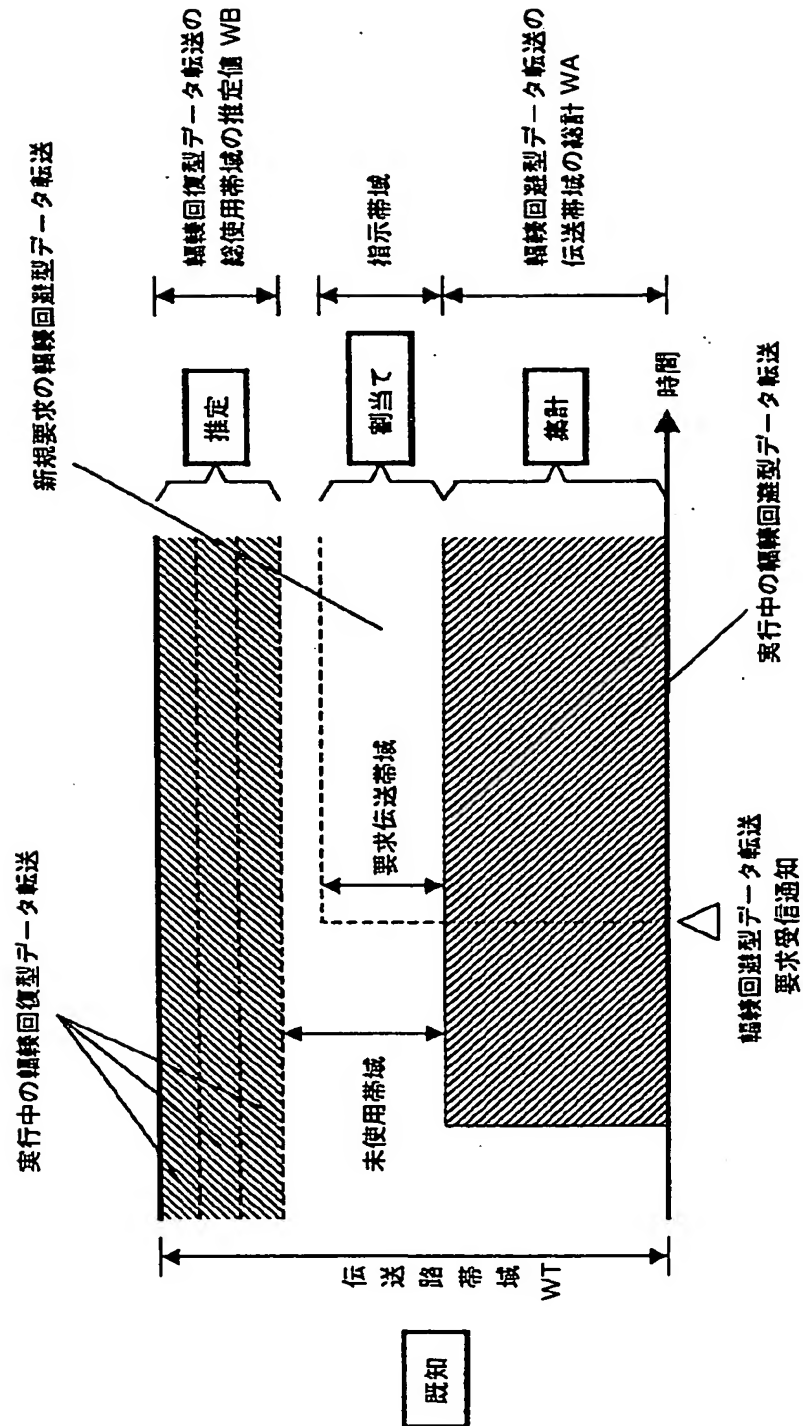


【図7】

本発明の伝送帯域推定手段の説明図

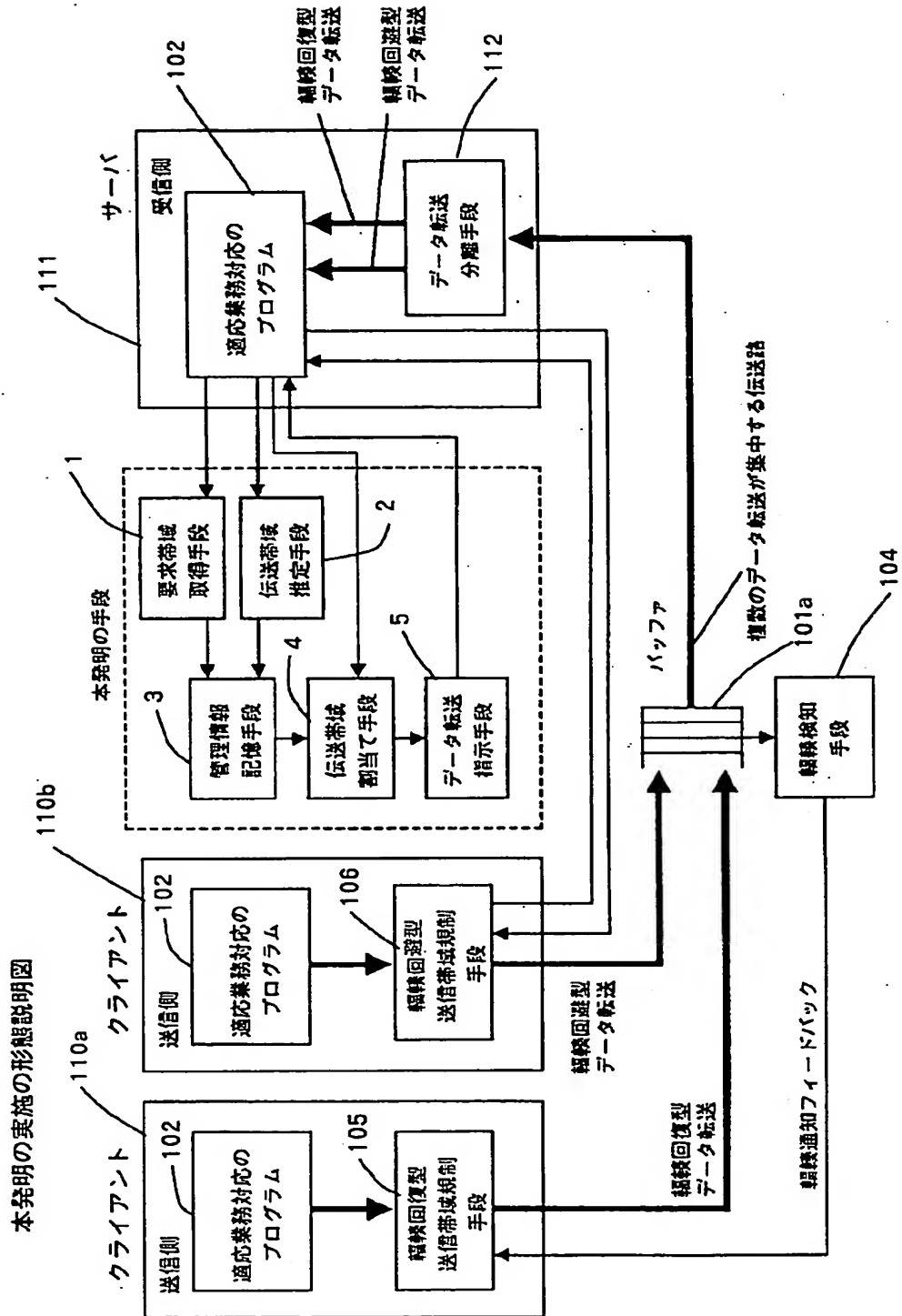


【図8】

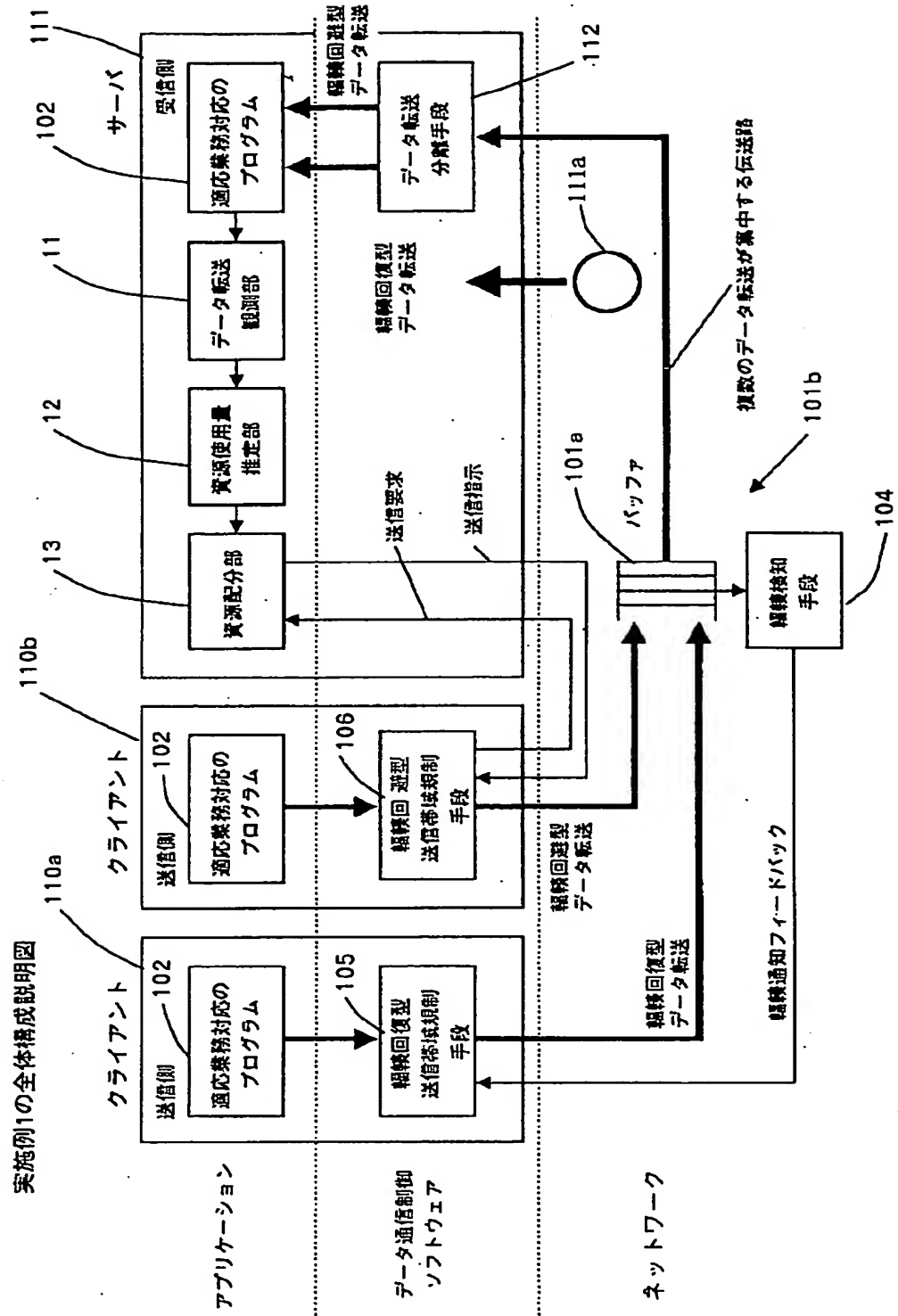


本発明の伝送帯域割当て方法の概略説明図

【図9】

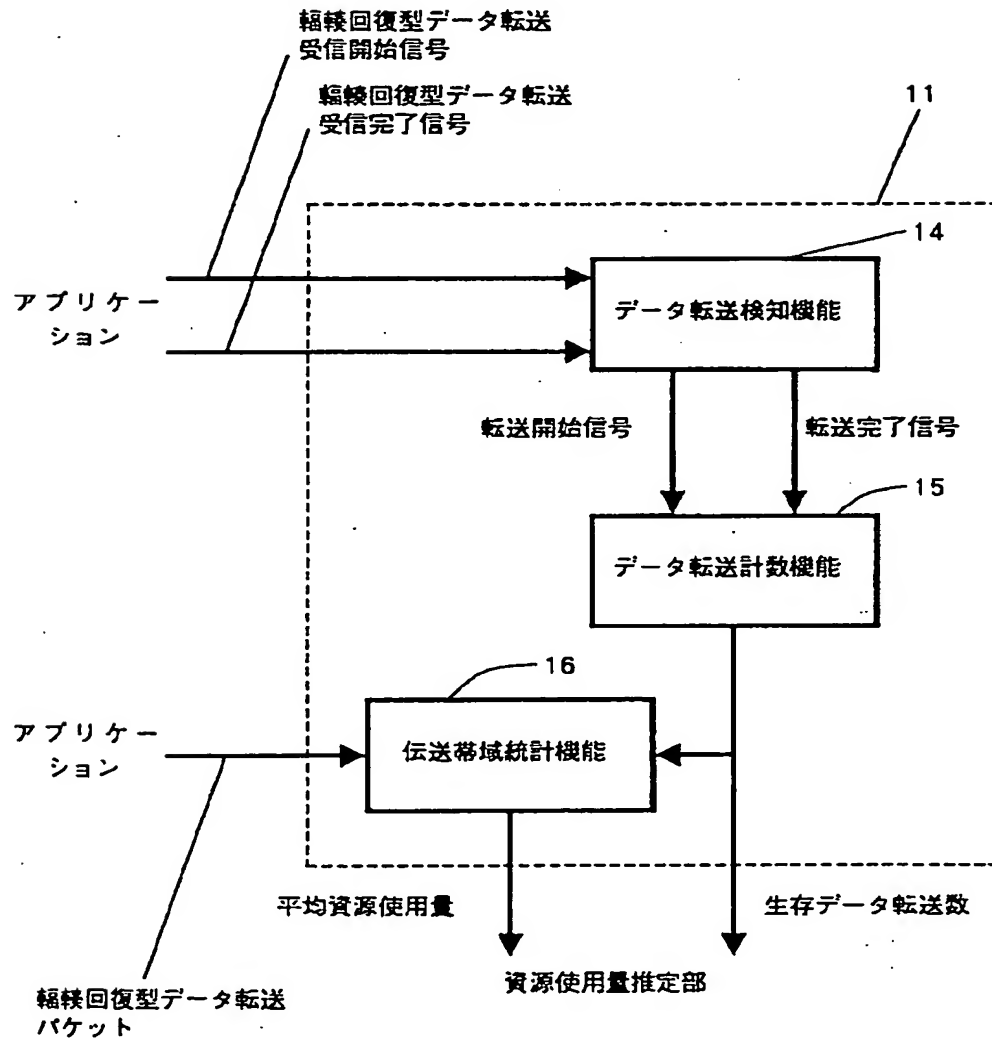


【図10】



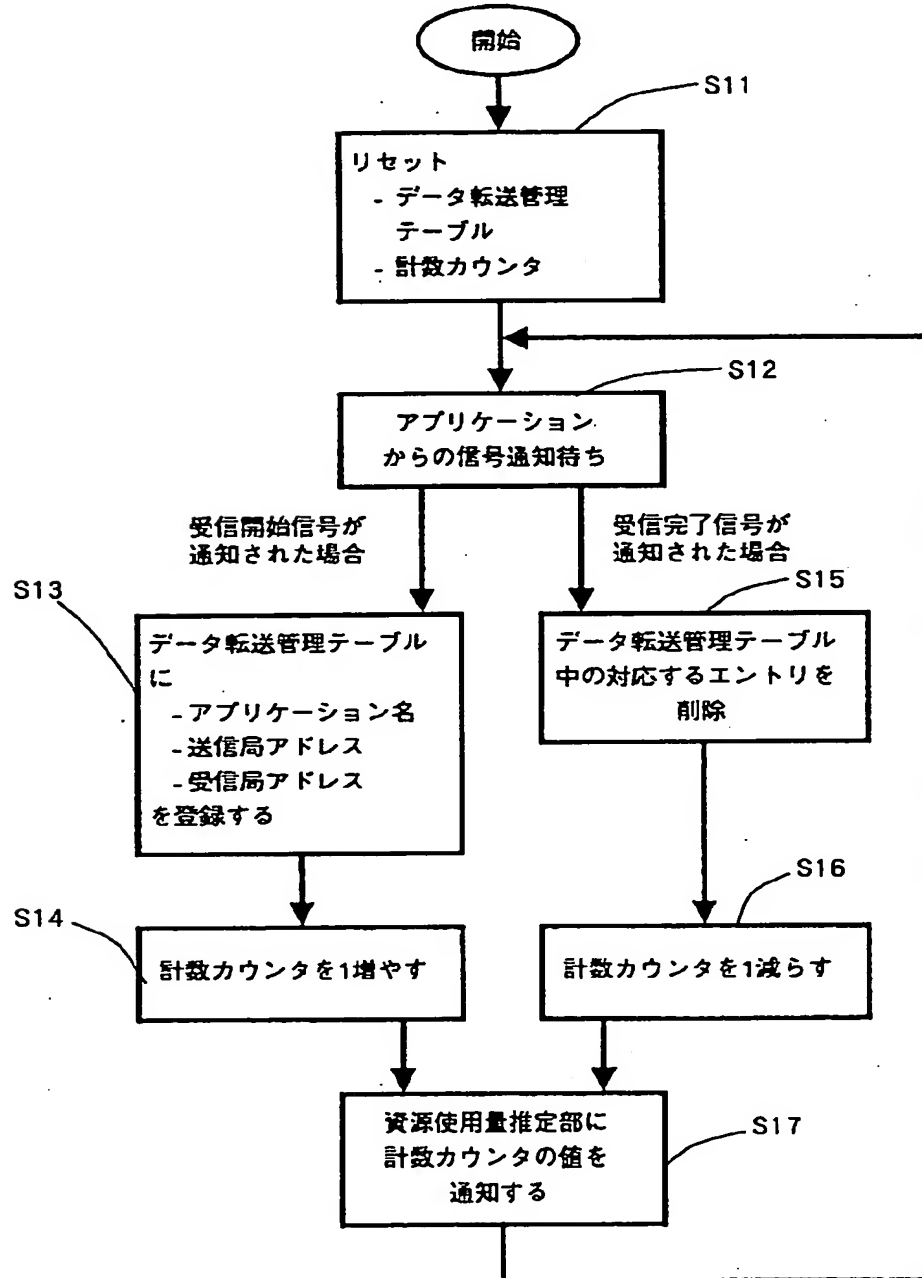
【図11】

実施例1のデータ転送観測部の構成説明図



【図12】

実施例1のデータ転送観測観測部の動作説明図



【図14】

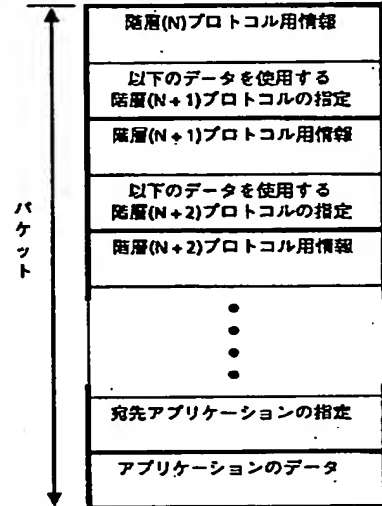
【図27】

階層型データ転送プロトコルによるパケット構造説明図

実施例1の伝送帯域統計テーブルの説明図

生存データ転送数 [本]	累計受信時間 [sec]	累計受信パケット数 [個]
1		
2		
3		
4		
5		

•
•
•



【図21】

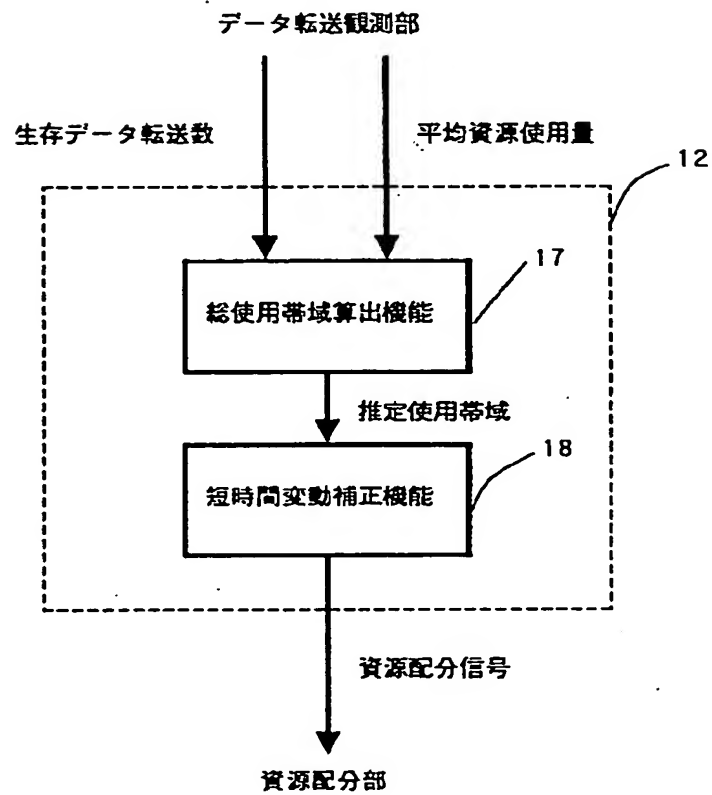
実施例1の資源使用量管理テーブルの説明図

	プロトコル識別子	送信局アドレス	受信局アドレス	送信帯域値 [Mbps]
1				
2				
3				

•
•
•

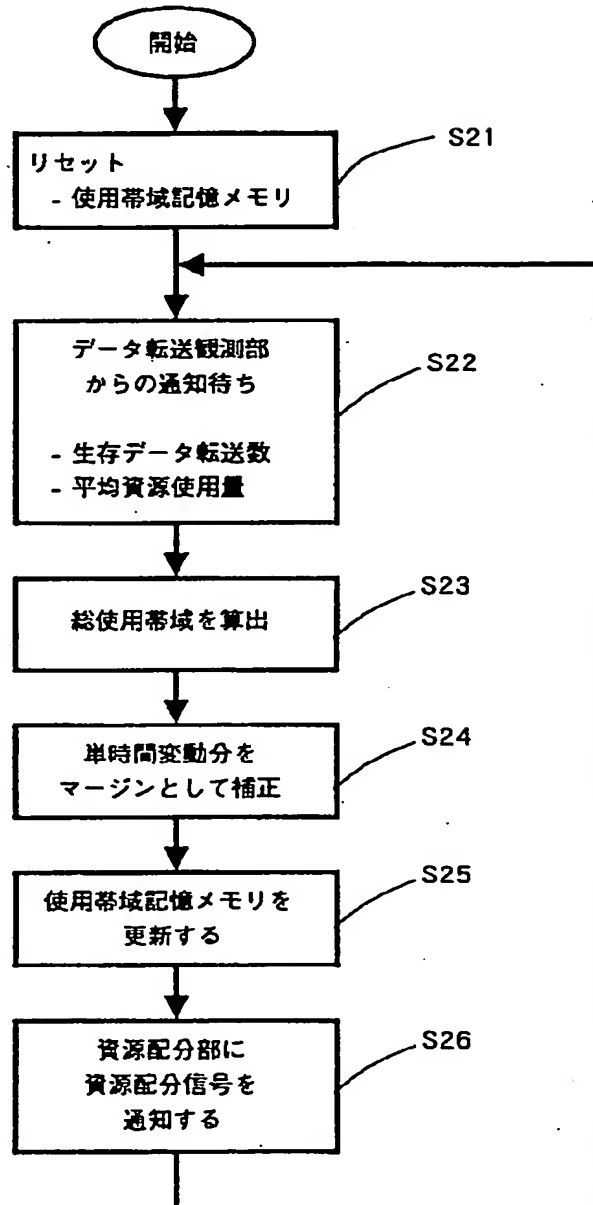
【図15】

実施例1の資源使用量推定部の構成説明図



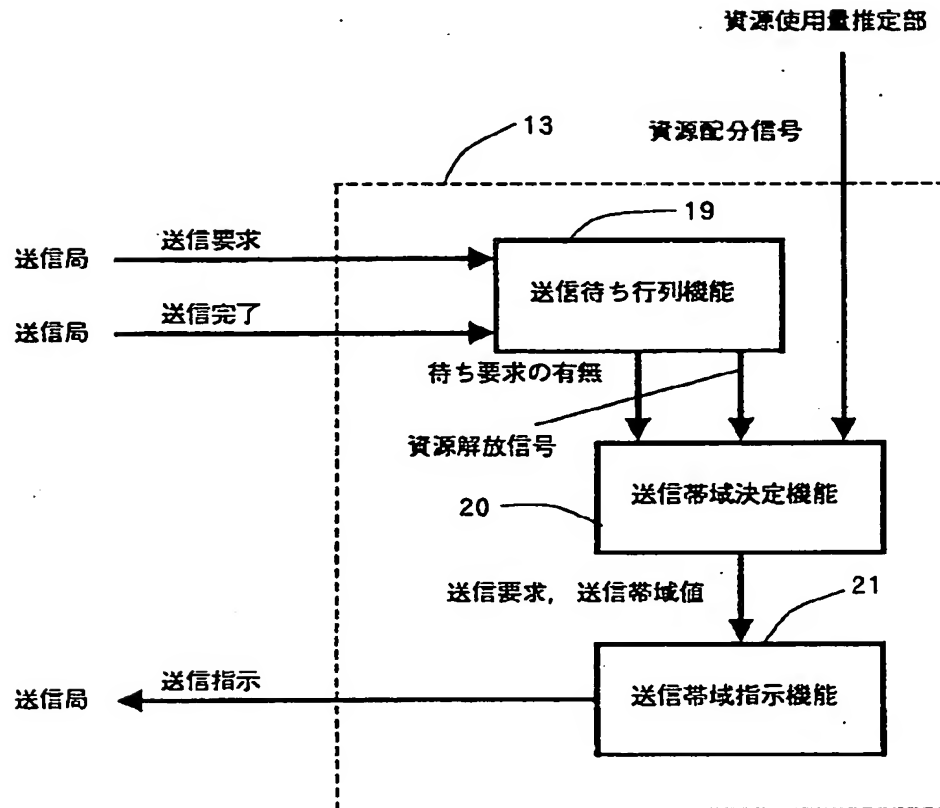
【図16】

実施例1の資源使用量推定部の動作説明図



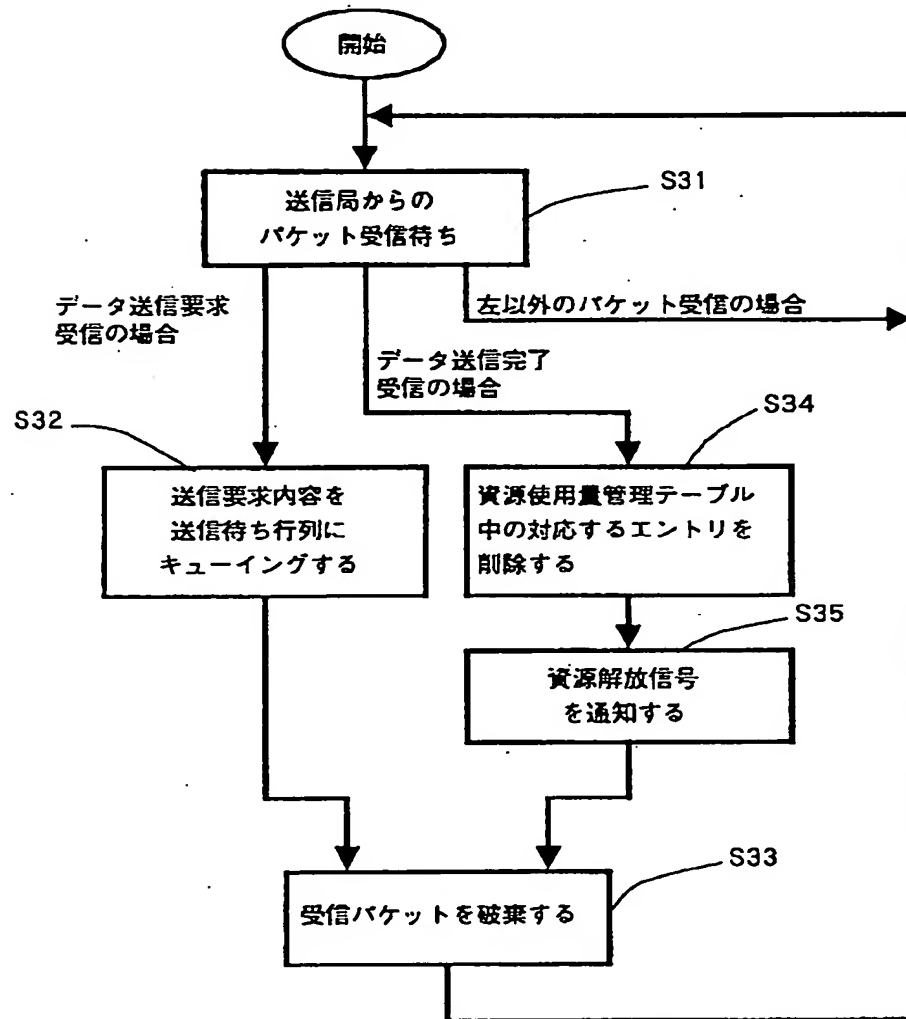
【図17】

実施例1の資源配分部の構成説明図



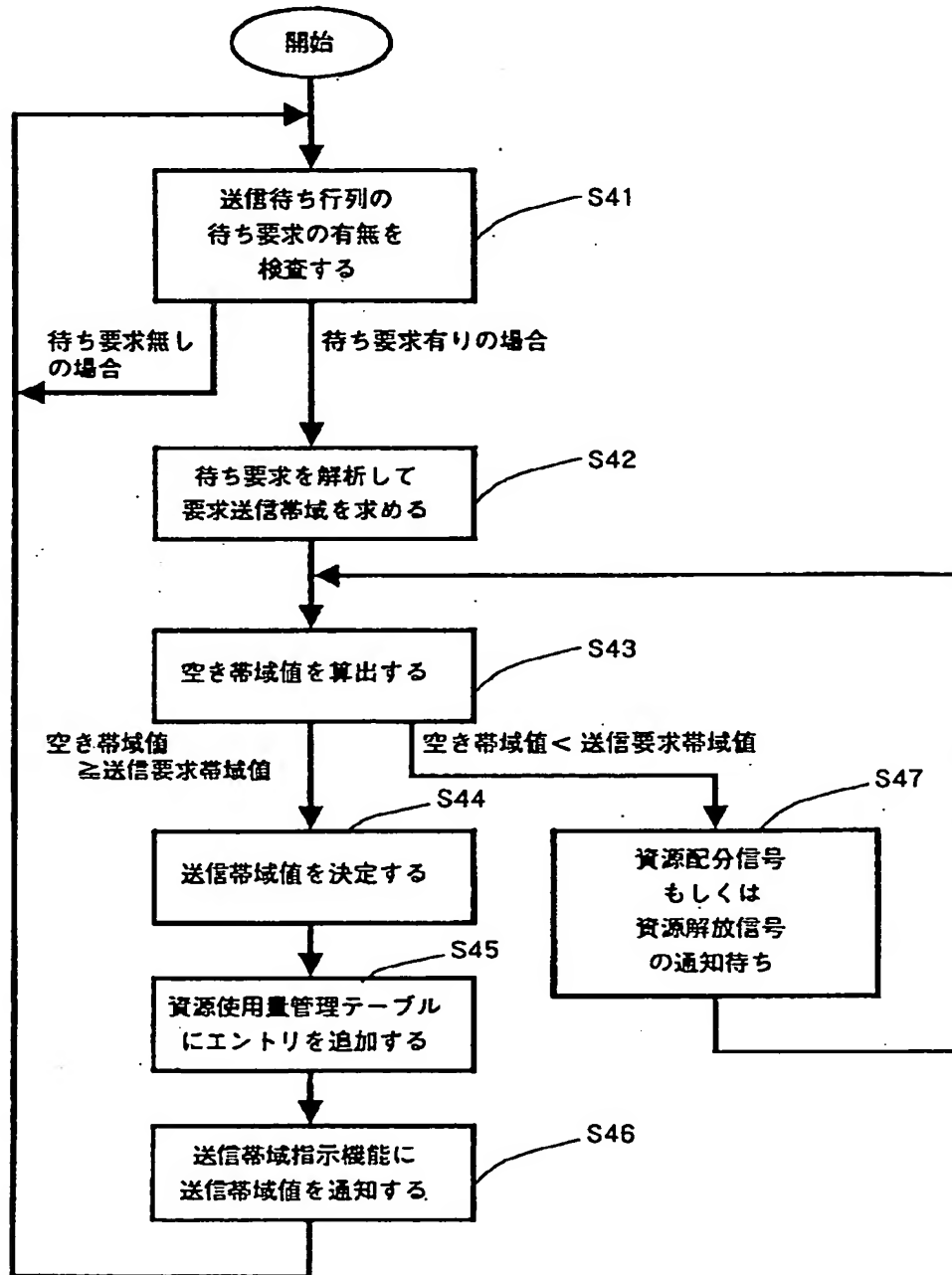
【図18】

実施例1の資源配分部の送信待ち行列機能の動作説明図

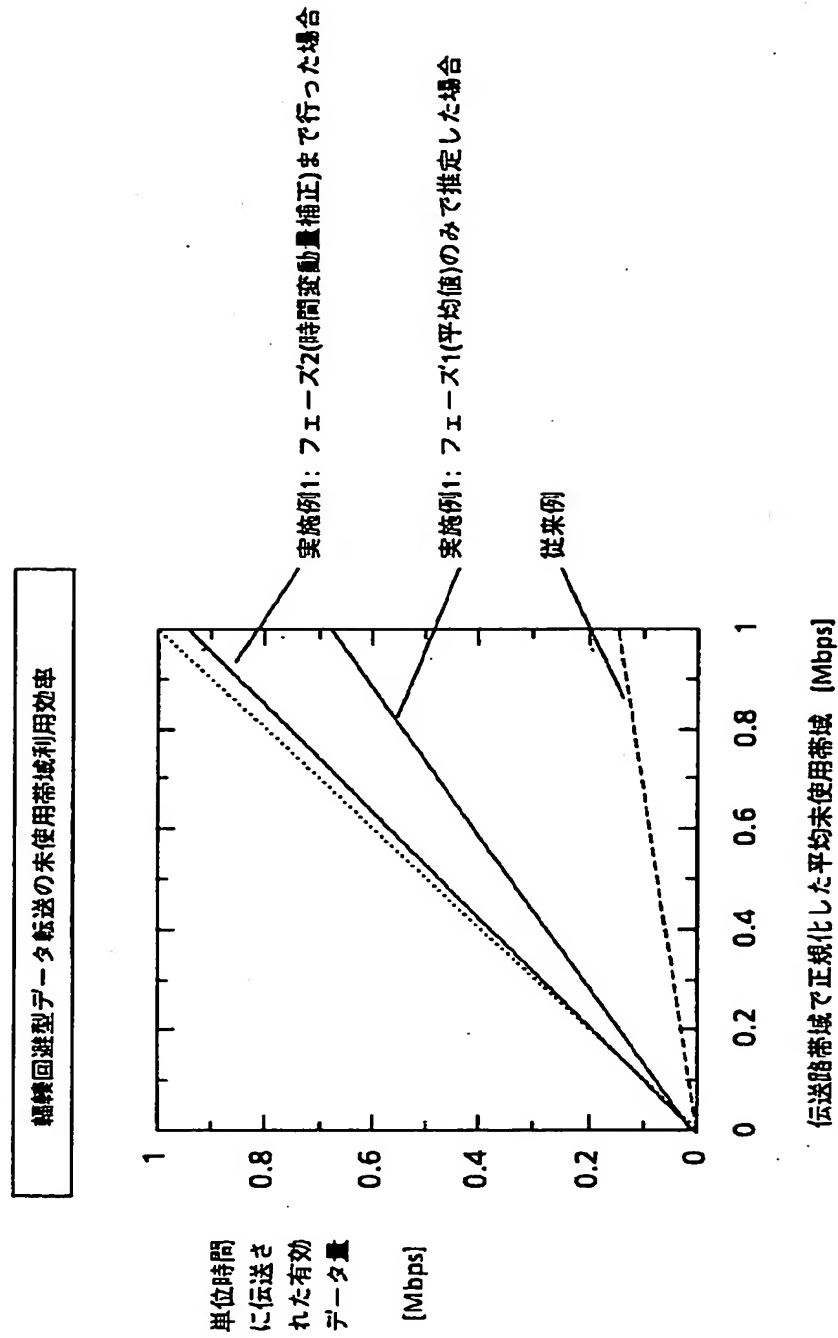


【図19】

実施例1の資源配分部の送信帯域決定機能の動作説明図

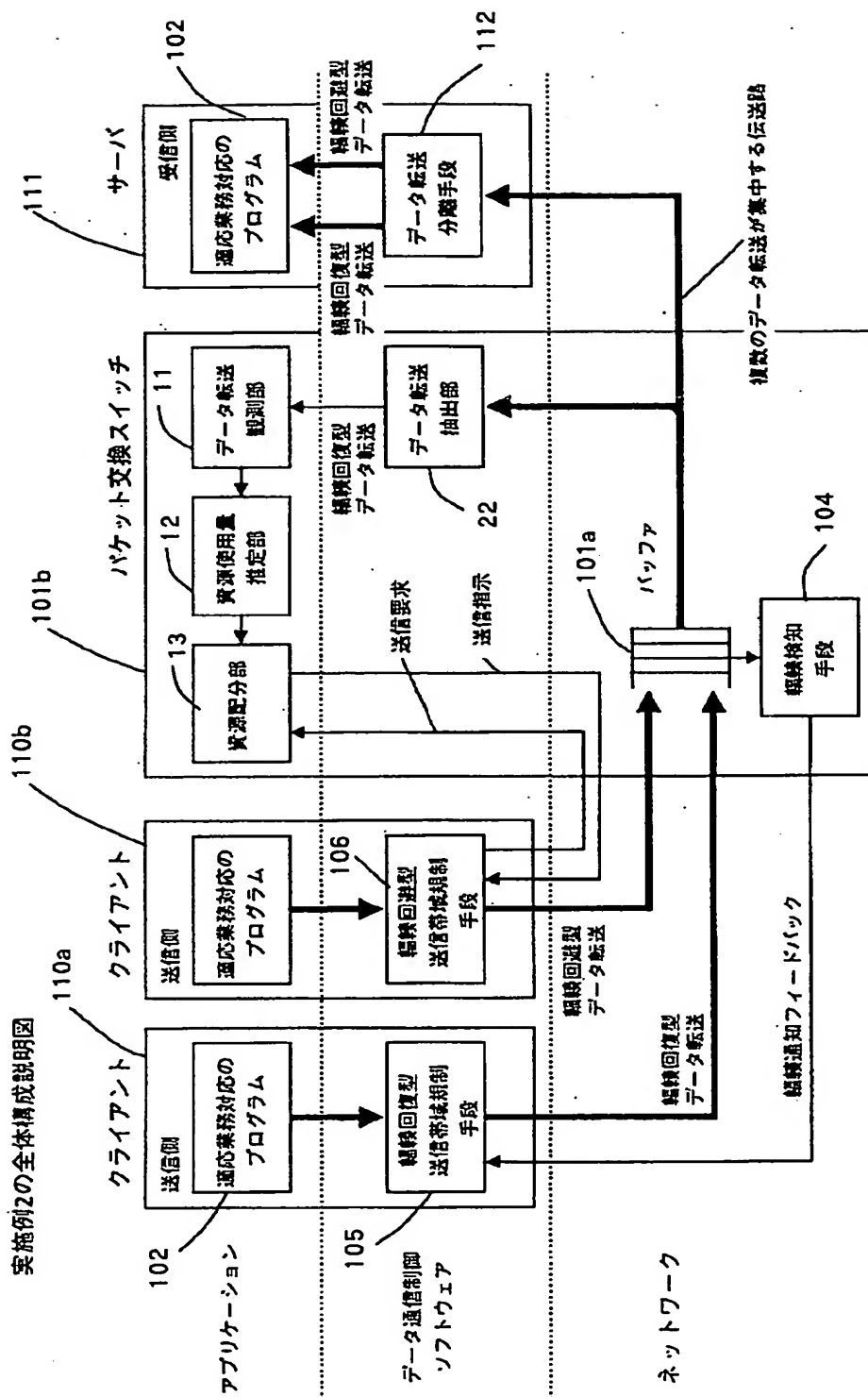


【図22】



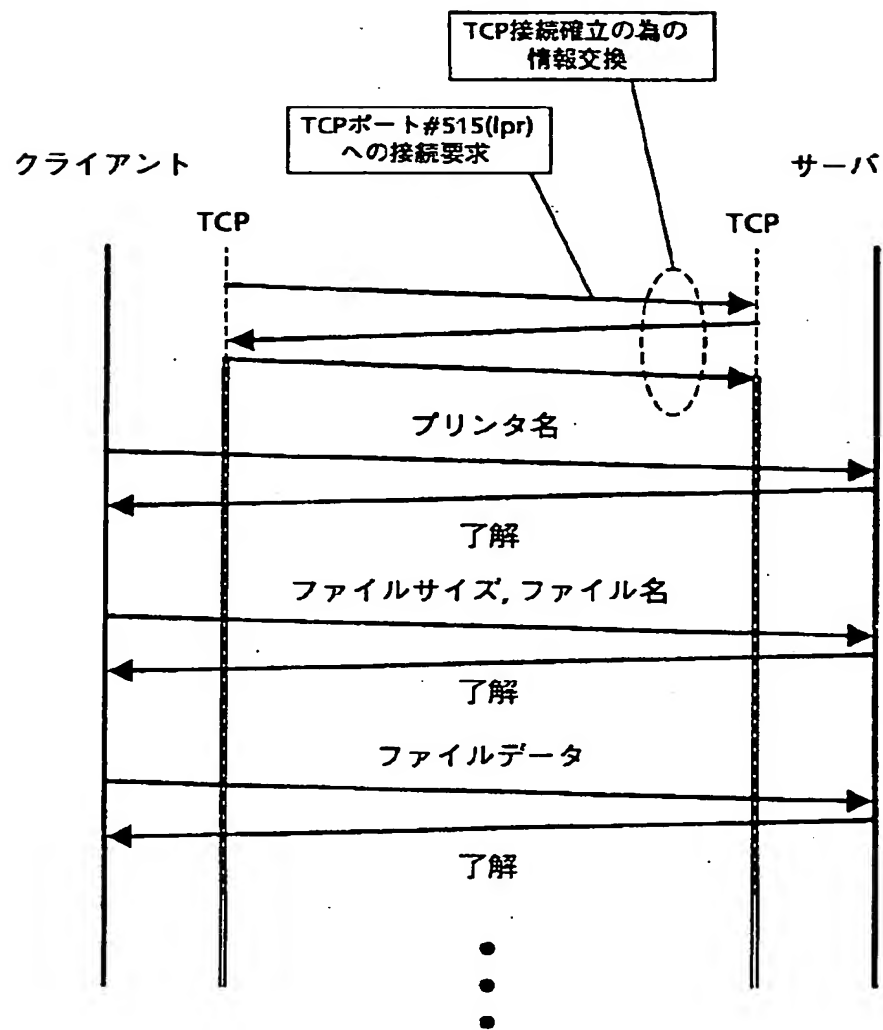
実施例の評価説明図

実施例2の全体構成説明図



【図24】

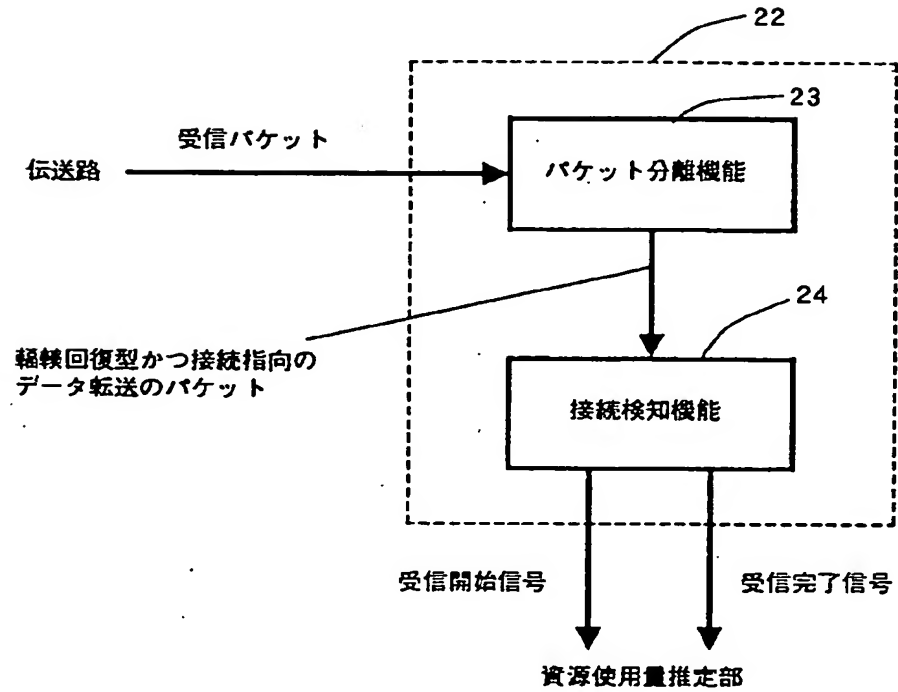
従来技術 TCP を用いた接続動作の例：lpr



lpr の場合、クライアントからサーバへのプリントジョブ転送に先立ち TCPポート505への接続が行われる。

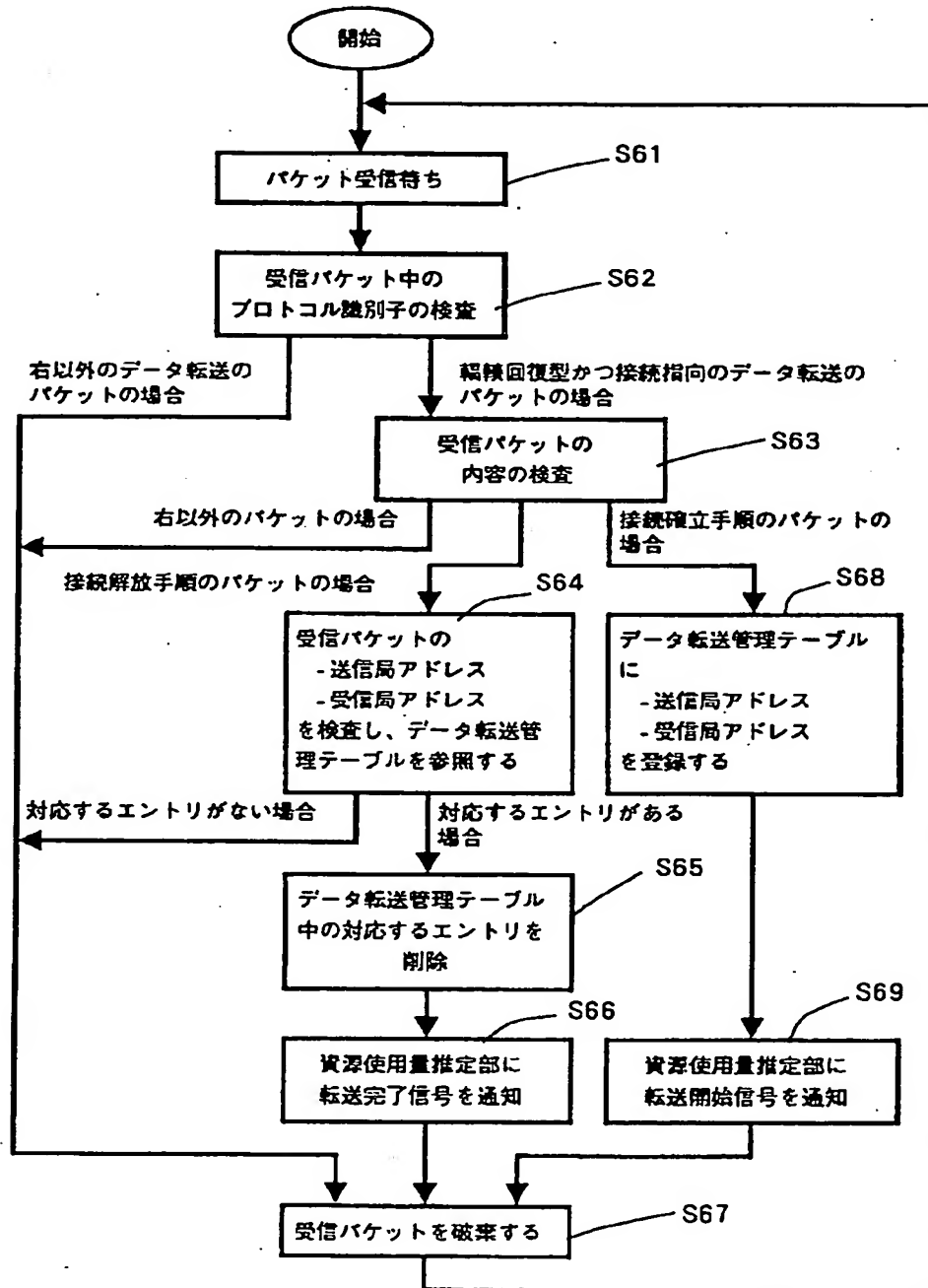
【図25】

実施例2のデータ転送抽出部の構成説明図



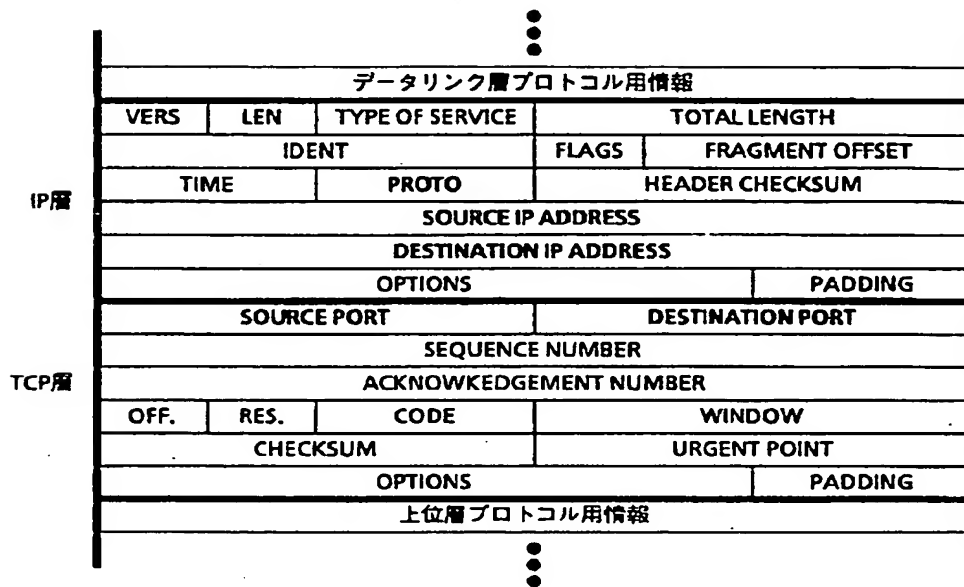
【図26】

実施例2のデータ転送観測観測部の動作説明図



【図28】

従来技術TCP/IPによるコネクション、および、プロトコル指定の説明図



IP層のフィールド)

PROTO: 上位層プロトコル識別子(ex. TCP = 6)
 SOURCE IP ADDRESS: 送信局IPアドレス
 DESTINATION IP ADDRESS: 受信局IPアドレス

TCP層のフィールド)

SOURCE PORT: 送信ポート番号
 DESTINATION PORT: 受信ポート番号

【図29】

実施例2のプロトコル情報テーブルの説明図

	プロトコル識別子	輻輳制御方式	接続形式	接続確立手順	接続解放手順
1					
2					
3					

•
•
•

輻輳制御方式

1: 回復型輻輳制御

0: 回避型輻輳制御

接続形式

1: 接続指向

0: それ以外

【図30】

実施例2のデータ転送管理テーブルの説明図

	プロトコル識別子	送信局アドレス	受信局アドレス	接続ポート番号
1				
2				
3				

•
•
•